APPLICATION

FOR

UNITED STATES LETTERS PATENT

TITLE: HEPADNAVIRUS PRE-S PROTEIN FRAGMENTS

APPLICANT: SHUPING TONG
JACK R. WANDS
JISU LI

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EL485680996US

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D C. 20231.

March 27, 2001
Date of Deposit

Samantha Bell
Signature

Samantha Bell
Typed or Printed Name of Person Signing Certificate

HEPADNAVIRUS PRE-S PROTEIN FRAGMENTS

Cross-Reference to Related Applications

This application is a divisional of USSN 09/361,707, filed July 27, 1999, now allowed, which is a divisional of USSN 08/683,262, filed July 18, 1996, now U.S. Patent No. 5,929,220, which claims benefit from USSN 60/001,371, filed July 21, 1995.

Statement as to Federally Sponsored Research

This invention was supported in part by NIH grants CA-35711, AA-08169, and AA-02666. The Government has certain rights to the invention.

Background of the Invention

The field of the invention is cellular receptors for viruses of the family Hepadnaviridae, and nucleic acids encoding the same.

The human hepatitis B virus (HBV) and related animal viruses that infect woodchucks, ground squirrels, Pekin ducks, and herons form a group of hepatotrophic DNA viruses in the family Hepadnaviridae. In humans, HBV causes acute and chronic liver disease and hepatocellular carcinoma.

The initial event of infection, interaction between the viral envelope protein and specific cellular receptor(s), is poorly understood. Hepadnaviruses express at least two co-terminal envelope proteins from a single envelope gene by alternative use of in-frame AUG codons. The large envelope protein (pre-S/S protein) of duck hepatitis B virus (DHBV) contains a 161-163 amino acid segment called the pre-S domain, and a carboxyterminal 167 amino acid segment called the S domain. The small envelope protein (S protein alone) is produced by translation from an internal AUG codon. The large envelope protein of HBV is similar, but has pre-S1 and pre-S2 domains in place of the single DHBV pre-S domain. As a result, two pre-S containing proteins are produced: a large envelope protein (preS1+preS2+S) and a middle envelope protein (preS2+S). The large envelope protein is myristylated and phosphorylated (Grgacic et al., J. Virol. 68:7344-7350, 1994; Macrae et al., Virology 181:359-363, 1991; Persing et al., J. Virol. 61:1672-1677, 1987). The large envelope protein mediates infection by DHBV and by

hepatitis delta virus (HDV), which borrows the envelope proteins of other hepadnaviruses to enter a hepatocyte (Fernholz et al., Virology 197:64-73, 1993; Summers et al., J. Virol. 65:1310-1317, 1991; Sureau et al., J. Virol. 67:366-372). The pre-S domain is believed to be responsible for binding a cellular receptor. Although several cellular proteins bind the HBV envelope, none have been shown to be the actual receptor (Budkowska et al., J. Virol. 67:4316-4322, 1993; Budkowska et al., J. Virol. 69:840-848, 1995; Hertogs et al., Virology 197:549-557, 1993; Mehdi et al., J. Virol. 68:2415-2424, 1994; Neurath et al., J. Exp. Med. 176:1561-1569, 1992; Pontisso et al., J. Gen. Virol. 73:2041-2045, 1992).

Since no cell culture system is available for the study of HBV, DHBV was developed as a model system. DHBV infection of ducklings and primary duck hepatocytes has been well characterized (Pugh et al., Virology 172:564-572, 1989; Tuttleman et al., J. Virol. 58:17-25, 1986).

Summary of the Invention

The invention features a purified nucleic acid that encodes a member of the hepadnavirus family of cellular receptors, or, where the receptor is a complex of two or more polypeptides, a component thereof. By "a member of the hepadnavirus family of cellular receptors" (hereafter a "hepadnavirus receptor") is meant a protein that binds the pre-S domain of the hepadnavirus large envelope protein so as to mediate or induce entry of a hepadnavirus virion into a host cell. A "hepadnavirus receptor", as used herein, can be the whole receptor, where the receptor is a monomer, or a subunit of a hepadnavirus receptor that binds the pre-S receptor binding site of the pre-S domain.

The identity of a hepadnavirus receptor that "mediates entry of the hepadnavirus virion into a host cell" can be confirmed using two biological activity assays. First, an antibody preparation specific for a member of the hepadnavirus family of receptor proteins should have the ability to block, inhibit, or reduce hepadnaviral infection of, or entry into, a cell, the virus being capable of infecting the same cell type in the absence of antibody. Either polyclonal or monoclonal receptor-specific antibodies can be used. For the second assay, a cDNA that encodes a member of the hepadnavirus family of receptor proteins is transfected into a cell line. The cell line is one that is ordinarily not a target cell for a hepadnavirus. Transfection of receptor-encoding cDNA sequence into these

cells should confer properties on the cell line that enable them to be infected by a hepadnavirus, or enable the virus to bind to the cell surface. By performing these two afore-mentioned assays, a hepadnavirus receptor is distinguished from a non-receptor hepadnavirus pre-S binding protein.

By "purified nucleic acid" is meant a nucleic acid that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (i.e., one at the 5' and one at the 3' end) in the naturally-occurring genome of the organism from which the nucleic acid of the invention is derived. The term encompasses deoxyribonucleic acid (DNA), for example, a cDNA or a genomic DNA fragment produced by the polymerase chain reaction (PCR), or produced by restriction endonuclease treatment. The cDNA or genomic DNA fragment can be incorporated into a vector, integrated into the genome of the same or a different species than the organism from which it was originally derived, linked to an additional coding sequence to form a hybrid gene encoding a chimeric polypeptide, or independent of any other nucleic acid sequences. The term also encompasses ribonucleic acid (RNA). The nucleic acid may be double-stranded or single-stranded, sense or antisense.

Examples of purified nucleic acids of the invention include those which encode amino acid sequences substantially the same as those shown in Figs. 18 and 19; and those having sequences that are either identical to, or hybridize under conditions of high or moderate stringency to, the 2.5 kb p170 cDNA included in the ATCC deposit Ep170pUC, designated No. 69869. High stringency conditions are herein defined as the following: hybridizing with 50% deionized formamide, 800 mM NaCl; 20 mM Pipes, pH 6.5, 0.5% SDS, 100 μg/ml denatured, sonicated salmon sperm DNA at 42°C for 12-20 hours, washing with 30 mM NaCl/3.0 mM sodium citrate (0.2 X SSC)/0.1% SDS at 55°C, while moderate stringency conditions are as follows: hybridizing with 50% deionized formamide, 800 mM NaCl; 20 mM Pipes, pH 6.5, 0.5% SDS, 100 μg/ml denatured, sonicated salmon sperm DNA at 42°C for 12-20 hours, washing with 75 mM NaCl/7.5 mM sodium citrate (0.5 X SSC)/0.1% SDS at 55°C. Such hybridization conditions are useful in a method of identifying a nucleic acid sequence encoding a hepadnavirus receptor polypeptide. The method involves providing a genomic or cDNA library; contacting the library with a nucleic acid that encodes a portion of a hepadnavirus

receptor, e.g., the duck p170 receptor, under conditions permitting hybridization between the nucleic acid and a homologous nucleotide sequence in the library; and identifying a clone from the library which hybridizes to the nucleic acid, hybridization being indicative of the presence in the clone of a nucleotide sequence homologous to a hepadnavirus receptor-encoding nucleotide sequence.

The invention also includes fragments of a purified nucleic acid that encodes a member of the hepadnavirus family of receptors. Examples include a nucleic acid of at least 20 nucleotides in length, or at least 30 or 50 nucleotides in length, that includes a strand which hybridizes under high stringency conditions to either the sense or antisense strand of a nucleic acid encoding part or all of a naturally occurring hepadnavirus receptor polypeptide. A nucleic acid fragment is useful, e.g., as a probe for identifying additional members of the hepadnavirus family of receptors, or for administering a portion of a hepadnavirus receptor sequence to a cell, e.g., a cell in a patient, using gene therapy techniques. Such portions of a hepadnavirus can include the peptides 1-4 shown in Fig. 14A.

The invention also includes vectors (e.g., plasmids, phage, or viral nucleic acid) or cells (prokaryotic or eukaryotic) which contain nucleic acids encoding any of the various hepadnavirus receptors of the invention. The vector can be any vector suitable for maintaining or making multiple copies of a nucleic acid of the invention, or can be one that is suitable for administering a nucleic acid of the invention to a cell or to a mammal infected with a hepadnavirus, e.g., to a human patient infected with HBV, or to cells removed from the patient for ex vivo gene therapy. Examples of vectors useful in a method of inhibiting hepadnavirus replication include, but are not limited to, adenovirus vectors, adeno-associated vectors, and retroviral vectors. Any of the various vectors of the invention can be included in a therapeutic composition along with a pharmaceutically acceptable carrier.

A purified nucleic acid of the invention can be under the transcriptional control of a heterologous promoter (i.e., a promoter other than one naturally associated with the given receptor gene of the invention). The promoter may direct the expression of the nucleic acid of the invention in a particular tissue or at a particular stage of development. A nucleic acid of the invention can also be in the form of a transgene in a transgenic non-

human animal, e.g., in a mouse. A transgenic animal bearing a transgene that encodes a hepadnavirus receptor is useful as an animal model to assay potential reagents for treating a hepadnavirus infection.

The invention also features hepadnavirus receptor polypeptides encoded by any of the various nucleic acids of the invention, for example, recombinant polypeptides expressed by a cell transformed with the nucleic acid. The polypeptide can be included in a therapeutic composition as an active ingredient, along with a pharmaceutically acceptable carrier, or it can be expressed from the nucleic acid within a target cell. The invention also includes an antibody that forms an immune complex with a hepadnavirus receptor of the invention. The antibody can be included in a therapeutic composition along with a pharmaceutically-acceptable carrier, or can be packaged in the form of a kit to be used as a diagnostic reagent. As a therapeutic, the antibody is useful for reducing the level of hepadnaviral infection in an animal, e.g., a human patient. The method involves administering a therapeutic preparation of the antibody to the animal in a dosage effective to inhibit the infection.

The invention also features a ligand that binds a member of the hepadnavirus family of receptors, hereafter referred to as a "p170 ligand." By "p170 ligand" is meant a molecule that specifically binds to the p170 binding site in the pre-S domain, in a mode that is competitive with the naturally-occurring hepadnavirus envelope. Preferably, a p170 ligand is a portion of the pre-S domain (hereafter a pre-S polypeptide) which binds the hepadnavirus receptor. The amino acid sequence of a pre-S polypeptide is substantially identical to a region of a wild type hepadnavirus pre-S domain (see, e.g., Figs. 16 and 17), the region being smaller in length than the complete amino acid sequence of the wild type pre-S domain, and being one that binds to a hepadnavirus receptor. A pre-S polypeptide of the invention is at least twelve amino acids, preferably at least nine amino acids, or more preferably at least six amino acids in length. By "binds to a hepadnavirus receptor" is meant that the p170 ligand forms a specific interaction with the receptor sufficient for copurification of the two components by an antibody specific for one of them. Alternatively, binding is indicated by co-purification on a GST affinity column, as described herein.

Where the hepadnavirus is DHBV, the aminoterminal amino acid of the pre-S p170 polypeptide amino acid sequence corresponds to a position selected from the group consisting of positions 1 to 87, or amino acids 25 to 87, 59 to 87, 70 to 87, or 80 to 87, all inclusive, of the amino acid sequence of Fig. 16. The carboxyterminal amino acid of the amino acid sequence can correspond to a position selected from the group consisting of positions 102 to 161, 104 to 161, 126 to 161, or 138 to 161, all inclusive, of the amino acid sequence of Fig. 16. Examples of pre-S polypeptides of the invention can include, but are not limited to, those including amino acids 25-161, 59-161, 71-161, 80-161, 87-161, 1-138, 1-126, 1-104, 1-102, 25-104, 25-102, 80-104, 25-126, 59-126, 71-126, 42-102, 59-104, all inclusive. Since applicants have identified the two basic amino acid residues at positions 95 and 97 as being essential for binding to the hepadnavirus receptor, it is recognized that a pre-S polypeptide encompassing amino acids 95 and 97 is a suitable p170 ligand. The full DHBV pre-S p170 nucleotide and amino acid sequences are shown in Fig. 16 (Mandart et al., J. Virol., 49:782-792, 1984).

Corresponding regions of the receptor binding site in the human HBV pre-S domain will be easily identified from the sequence alignment shown in Fig. 20, and by routine techniques of sequence analysis. For example, the receptor binding site in the HBV pre-S domain includes amino acids 89-104. The aminoterminal amino acid of the binding site preferably corresponds to one of the amino acids between positions 71 and 89, inclusive, of the amino acid sequence of Fig. 20. The carboxyterminal amino acid of the receptor binding site preferably corresponds to one of the amino acids between positions 104 to 118, inclusive, of Fig. 20. Examples of HBV pre-S polypeptides of the invention can include, but are not limited to, those including amino acids 89-104, or 71-118, of Fig. 20. Since applicants have identified the arginine residue at position 99 as being essential for binding to the hepadnavirus receptor, it is recognized that a pre-S polypeptide encompassing amino acid 99 is a suitable hepadnavirus receptor ligand. The full HBV pre-S nucleotide and amino acid sequences of several strains that infect primates are shown in Fig. 17 (Norder et al., Virology, 198:489-503, 1994). Using pre-S protein fused to glutathione S-transferase and immobilized on Sepharose beads, we have now identified an additional binding protein of 120 kDa (p120). p120 expression is restricted to the liver, kidney and the pancreas, the three major organs of DHBV

replication. While optimal p170 binding requires intact pre-S protein, binding to p120 occurs much more efficiently with a few N- or C-terminally truncated forms. The p120 binding site was mapped to residues 98-102 of the pre-S region, which overlaps with a cluster of known virus-neutralizing epitopes. Site-directed mutagenesis revealed residues 100-102 (Phe-Arg-Arg) as the critical p120 contact site; non-conservative substitution in any of the three positions abolished p120 binding. Double mutations at positions 100-102 markedly reduced DHBV infectivity in cell culture. Short pre-S peptides covering the clustered neutralizing epitopes (also p170/p120 binding sites) reduced DHBV infectivity in primary duck hepatocyte cultures. Thus, p120 represents a candidate component of the DHBV receptor complex.

The hepadnavirus receptors and pre-S polypeptides of the invention can be fused to a glutathione-S-transferase amino acid sequence (Smith et al., Gene 67:31, 1988). The polypeptide can be glycosylated or unglycosylated. "Glycosylated", as used herein, refers to having one or more covalently-linked carbohydrate moieties attached to the protein. By "unglycosylated" is meant lacking covalently-linked carbohydrate moieties. The hepadnavirus receptor can also be myristylated or unmyristylated, or phosphorylated or unphosphorylated, meaning that the hepadnavirus receptor or pre-S polypeptide has one or more covalently-attached myristic acid or phosphate groups, respectively.

The invention also includes a vaccine for the prevention of a hepadnaviral infection. A vaccine can be in the form of either an immunologically cross-reactive form of a naturally-occurring hepadnavirus receptor, or a nucleic acid encoding the same. A vaccine of the invention can also be in the form of a pre-S polypeptide, or a nucleic acid encoding a pre-S polypeptide. Where the vaccine is administered as a live cell vaccine, it may be desirable to inactivate the ability of the pre-S polypeptide to activate the receptor. Thus, a codon corresponding to an amino acid residue of the naturally occurring pre-S polypeptide can be deleted or altered to encode an amino acid residue different from the amino acid residue of the naturally occurring pre-S polypeptide; for the p170 ligand, this is preferably an amino acid residue selected from the group consisting of amino acids 95 and 97. Examples of suitable mutations of p170 include, but are not limited to, a substitution of Arg to Leu at position 95 (R95L), a substitution of Lys to Ser at position 95 (K95S), or a substitution of Arg to Cys at position 97 (R97C) of the sequence of Fig.

16. Codons encoding the hepadnavirus receptor binding-site of p170 can also be deleted from the vaccine, e.g., by a deletion of all of the codons encoding binding site amino acids, e.g., amino acids 87 to 102, or by deleting codons encoding amino acids 95 and 97, from the nucleic acids, thereby to delete receptor-binding capacity from the polypeptide it encodes. Corresponding modifications of p120 can be made and used in vaccines. Such polypeptide and nucleic acid-based vaccines are useful in a method of immunizing an animal against hepadnaviral infection, by introducing an immunizing amount of the nucleic acid or polypeptide into the animal.

Also encompassed within the invention is a method of producing a hepadnavirus receptor polypeptide, by the steps of (a) providing the cell that includes a nucleic acid encoding a hepadnavirus receptor; and (b) culturing the cell under conditions permitting expression of the polypeptide from the nucleic acid.

Finally, the invention includes a method for identifying an antagonist to a hepadnavirus receptor. The method involves (a) contacting a hepadnaviral receptor, in the presence and in the absence of the candidate antagonist, with a form of hepadnavirus envelope protein; and (b) comparing the level of binding of the receptor to the form of hepadnavirus envelope protein in the presence of the candidate antagonist, with the level of binding of the receptor to the form of hepadnavirus envelope protein in the absence of the candidate antagonist. A lower level of binding in the presence of the candidate antagonist than in its absence indicates that the candidate antagonist is capable of competing with the form of hepadnavirus envelope protein for binding to the receptor. "Antagonist", as used herein, refers to a chemical substance that inhibits an activity of the receptor, such as its ability to bind a ligand or agonist, e.g., a hepadnavirus. A "form of hepadnavirus envelope protein" can be a naturally occurring hepadnavirus; a hepadnavirus envelope particle; a hepadnavirus subparticle; an envelope protein; a hepadnavirus pre-S protein; or a pre-S polypeptide of at least six amino acids that includes the hepadnavirus receptor-binding domain as characterized herein. The hepadnaviral receptor polypeptide can be provided in the form of a cultured eukaryotic cell transfected with a nucleic acid that encodes a hepadnavirus receptor, the receptor being expressed in the cell preferably as a cell surface receptor. The hepadnaviral receptor can also be provided in the form of a transgenic non-human animal bearing a

transgene, the transgene including a nucleic acid that encodes a hepadnavirus receptor as a source of the receptor.

As used herein, a "hepadnavirus" refers to a member of the Hepadnaviridae family of viruses, including, but not limited to, hepatitis B virus and hepatitis delta virus (Wang et al., Nature, 323:508-13, 1986). Cellular receptors that interact with other hepadnavirus species are included within the scope of the invention. Examples include, but are not limited to, avian strains such as duck hepatitis B virus (DHBV; Mandart et al., J. Virol. 49:782-792, 1984; Mason et al. J. Virol. 36:829-36, 1978), or heron HBV (Sprengel et al., J. Virol., 62:3832-39, 1988); woodchuck hepatitis virus (WHV; Summers et al. Proc. Natl. Acad Sci. USA, 75:4533-37, 1978), and squirrel hepatitis virus (e.g., Marion et al. Proc. Natl. Acad Sci. USA, 77:2941-45, 1980). These species can be useful laboratory models of the human hepatitis B virus. Examples of other hepadnaviruses within the scope of the invention include, but are not limited to, HBV strains infecting various human organs, including liver cells, exocrine and endocrine cells, tubular epithelium of the kidney, spleen cells, leukocytes, lymphocytes, e.g., splenic, peripheral blood, B or T lymphocytes, and cells of the lymph nodes and pancreas (see, e.g., Mason et al., Hepatology, 9:635-645, 1989). Various HBV strains within the scope of the invention include those disclosed by Norder et al., J. Virol., 198:489-503, 1994). The invention also applies to hepadnaviruses infecting non-human mammalian species, such as domesticated livestock or household pets. Nucleic acids derived from any of these species are useful for identifying and isolating further members of the hepadnavirus family of cellular receptors.

Where the method of inhibiting hepadnavirus replication is used to treat a hepadnaviral infection in an animal, a "naturally-occurring" hepadnavirus refers to a form or sequence of the virus as it exists in an animal, e.g., a natural isolate derived from an infected animal. In all other contexts, a "naturally-occurring" hepadnavirus is intended to be synonymous with the sequence known to those skilled in the art as the "wild type" sequence, e.g., the wild type pre-S protein sequences shown in Figs. 16 and 17. If an amino acid sequence of a pre-S protein of a hepadnavirus that is derived from a natural isolate differs from the conventionally accepted "wild type" sequence, it is understood that the sequence of the natural isolate may be the proper comparison sequence for

designing mutant polypeptides of the invention. The sequence of the natural isolate can be compared to the sequences cited herein to identify a receptor binding domain analogous to that of the DHBV pre-S domain.

Other terms and definitions used herein will be understood by those of routine skill in the art. For example, by "apparent molecular weight" is meant the molecular weight determined on a denaturing polyacrylamide gel by comparison with standards, e.g., protein standards, of known molecular weight. "Receptor", as used herein, refers to a molecule on the surface of a target cell that binds to and permits entry of its target ligand into the cell, e.g., a ligand such as a hepadnavirus virion. By "inhibiting hepadnavirus replication" is meant lowering the rate or extent of replication relative to replication in the absence of a mutant polypeptide of the invention. The term "fragment", as applied to a polypeptide will ordinarily be at least about 6 contiguous amino acids, typically at least about 9 or 12 contiguous amino acids, more typically at least about 20, or preferably at least about 30 or 35 or more contiguous amino acids in length. The term "fragment" as applied to a nucleic acid will ordinarily be at least about 15, typically at least about 20 or 30, more typically at least about 50, bp in length.

The methods, nucleic acids, and polypeptides of the invention can be used to inhibit entry of a hepadnavirus into a host cell, e.g., a cell of a mammal, e.g., a human patient, as an effective therapy for treating individuals with a persistent HBV infection, or as a means of reducing the risk of hepatocellular carcinoma in an infected animal. Polypeptides of the invention can be administered to an infected animal either directly or by gene therapy techniques. The screening methods of the invention are simple, rapid, and efficient assays designed to identify candidate antagonists or ligands, e.g., pre-S polypeptide ligands, with anti-hepadnaviral activity.

Other features and advantages of the invention will be apparent from the following detailed description and from the claims.

Brief Description of the Drawings

Fig. 1A is a Coomassie-blue stained SDS-polyacrylamide gel showing the migration of Glutathione-S-transferase (GST) fusion proteins. The molecular size of the protein markers are shown to the left. Fig. 1B is an autoradiographic image of a SDS-

polyacrylamide gel showing proteins isolated from metabolically labeled lysates of primary duck hepatocytes using GST fusion proteins as probes. Fig. 1C is an autoradiographic image of a SDS-polyacrylamide gel showing the detection of p170 on the cell surface. For Fig. 1B and Fig. 1C, the positions of p170 and protein size markers are shown to the right.

Fig. 2 is an autoradiographic image of cell surface proteins run on a SDS-polyacrylamide gel, comparing the specificity of labeling with lactoperoxidase versus the specificity of labeling with the iodogen method.

Fig. 3A is an autoradiographic image of an SDS-polyacrylamide gel showing competition of p170 binding by added DHBV particles. Fig. 3B is an autoradiographic image of an SDS-polyacrylamide gel showing reduced expression or inaccessibility of p170 in DHBV-infected hepatocytes.

Fig. 4 is an autoradiographic image of an SDS-polyacrylamide gel showing a reduced level of p170 when detected by a GST-pre-S protein in DHBV-infected liver tissue.

Fig. 5A is an autoradiographic image of an SDS-polyacrylamide gel showing the elution profile of p170 through a DEAE-cellulose column. Fig. 5B is an autoradiographic image of an SDS-polyacrylamide gel showing the tissue distribution of p170 in various duck tissues.

Fig. 6 is a schematic representation of pre-S deletion mutants and their binding results.

Fig. 7A is an autoradiographic image of an SDS-polyacrylamide gel showing the expression and purification of GST-pre-S deletion mutants. Molecular size markers are shown to the right. Fig. 7B is an autoradiographic image of an SDS-polyacrylamide gel showing the binding of p170 to GST-pre-S deletion mutants.

Fig. 8 is a schematic illustration that the p170 binding site coincides with a neutralizing epitope of the pre-S domain.

Fig. 9 is a schematic illustration showing which residues of the pre-S domain are required for p170 binding.

Fig. 10 is an autoradiographic image of a SDS-polyacrylamide gel showing the retention of p170 by different mutants.

Fig. 11 is an autoradiographic image of a SDS-polyacrylamide gel showing inhibition of DHBV infection by a pre-S polypeptide containing a p170 binding site.

Fig. 12 is a chart showing the location and nature of amino acid substitutions in mutants of the DHBV pre-S domain (SEQ ID NO:18).

Fig. 13A is an autoradiographic image of a SDS-polyacrylamide gel showing expression of GST-pre-S fusion proteins from pre-S mutant constructs. Fig. 13B is an autoradiographic image of SDS-polyacrylamide gel showing the binding capacity of pre-S mutants for p170.

Fig. 14A is an illustration of four peptide sequences from p170 (SEQ ID NOs:19-22). Fig. 14B shows the sequence similarity of peptides 1 (SEQ ID NO:19) and 2 (SEQ ID NO:20) to various carboxypeptidases.

Fig. 15 is an illustration of cloning p170 DNA from the p170 peptide sequences (SEQ ID NOs:23-32).

Fig. 16 is an illustration of the nucleic acid sequence (SEQ ID NO:33) of the DHBV pre-S gene and the corresponding translated amino acid sequence (SEQ ID NO:34).

Fig. 17 is an illustration of the nucleic acid sequence of the HBV pre-S gene and the corresponding translated amino acid sequence (SEQ ID NOs:35-62).

Fig. 18 is an illustration of the nucleic acid sequence (SEQ ID NO:63) of the 5' 1.1 kb of the p170 cDNA and the corresponding translated amino acid sequence (SEQ ID NO:64).

Fig. 19 is an illustration of the nucleic acid sequence (SEQ ID NO:65) of the 3' 460 bp of the p170 cDNA and the corresponding translated amino acid sequence (SEQ ID NO:66).

Fig. 20 is an illustration of the pre-S amino acid sequence in human HBV (SEQ ID NO:67) that corresponds to the p170 binding site in the DHBV pre-S protein (SEQ ID NO:68).

Fig. 21A and B is a pair of panels illustrating the recognition of a 120 kDa duck hepatocyte protein by several truncated forms of DHBV pre-S protein. The pre-S part of DHBV large envelope protein and its truncated forms (N-, C-, double-deletions) were expressed as GST fusion proteins and purified on glutathione-Sepharose beads. They

were incubated at 40°C with ^{35}S labeled primary duck hepatocyte lysates which had been precleared with Sepharose beads. After extensive washing with the lysis buffer, bound proteins were fractionated by reducing SDS 8% PAGE gel. ^{35}S labeled proteins were revealed by fluorography. A) Schematic representation of the pre-S constructs and their affinities for p120. B) Fluorograph. Molecular markers are shown at left and positions of p170 and p120 indicated.

Fig. 22 is a schematic diagram illustrating co-localization of p120/p170 binding sites and clustered neutralizing epitopes. The pre-S region is schematically shown at the top (amino acids 1-161). p170/p120 binding sites and the epitopes recognized by neutralizing (black bar) and non-neutralizing (white bar) mAbs are indicated. The epitopes 58-66, 91-99, 127-138, and 139-145 are according to Yuasa et al. (1991) Virology 181: 14-21; 83-90 and 100-107 are according to Chassot et al. infra. A nonessential region for viral infectivity is shown as a dotted bar.

Fig. 23 is a pair of Northern blots illustrating p120 retention by the intact pre-S construct. DHBV infected (left panel) or non-infected (right panel) duck liver was homogenized in lysis buffer. The cell lysate was incubated with intact pre-S construct 1-161 (lane 1) or a truncation construct 80-102 (lane 2). The proteins retained on sepharose beads were separated on SDS 8%PAGE gel and blotted onto nitrocellulose filter. After incubation with a rabbit anti-p170 antibody (upper panel) (16a) or anti-p120 antibody (lower panel), the protein bands were revealed by HRP-conjugated anti-rabbit serum and DAB. The positions of p170 and p120 are indicated.

Fig. 24 is a gel showing that p120 retention by the C-terminal deletion constructs of pre-S protein requires an exact truncation at residue 102 or 103. The five constructs used had common N-terminus at residue 80 but different C-terminus at residues 104, 103, 102, 101, and 100, respectively. Detection of p120 from ^{35}S labeled primary duck hepatocyte lysates was performed as described in the legend to Figure 21. Positions of p170 and p120 are indicated.

Fig. 25A and B are a table and a fluorograph illustrating the critical role of Pre-S residues 100-102 in p120 binding. Single amino acid substitutions were introduced into either construct 80-102 or construct 92-161 and effects on p120 binding examined with ^{35}S labeled primary duck hepatocyte lysates. A) Schematic representation of the mutants

and binding results. B) Fluorograph showing binding results of mutants in the construct 80-102 (left) or 92-161 (right). WT: wild-type (SEQ ID NO:69). Position of p120 band is indicated.

Fig. 26 is a gel showing detection of p120 in lysates of cell-surface biotinylated primary duck hepatocytes. Primary duck hepatocytes cultured for two days in Petri dishes were labeled with sulfo-LC-biotin. Cells were washed three times with PBS before and after labeling. The cell lysates were incubated with the following pre-S constructs: 80-102 (1); 80-104 (2); 25-102 (3); 1-161 (4). Positions of p120 and p170 are indicated. As a negative control, 6 mg of either surface (lane 5) or total biotinylated liver proteins (lane 6) was immunoprecipitated with a mAb M3A5 which recognizes an epitope shared by Golgi ß-COP, a Golgi membrane protein and microtubule-associated protein (MAP). The position of MAP doublet is indicated.

Fig. 27 is a gel showing impairment of DHBV infectivity by mutations at p120 binding site. Various DHBV mutants in an overlength DHBV genome were transfected into LMH cells in duplicate. DHBV particles were concentrated from pooled medium and equal amounts of virion particles were used to infect primary duck hepatocyte cultures for 6.5 hrs. Cells were harvested at day 7 post-infection. A) Southern blot analysis of viral particles secreted to LMH culture medium at day 3 post-transfection. B) Southern blot analysis of another aliquot of day 3 viral particles pre-treated with Pronase and DNase I. C) DHBV DNA associated with intracellular core particles at day 7 post-transfection. D) Intracellular DHBV DNA in primary duck hepatocytes infected for 6.5 hr with virus particles produced in LMH cells. Lane 1: R101I/R102D; 2: R101L/R102L; 3: F100V/R101L; 4: Y103C/Q104F; 5: K95S/R97L/E98A; 6: wild-type.

Fig. 28 is a Southern blot illustrating the inhibition of DHBV infection of primary duck hepatocytes by short pre-S peptides. The pre-S peptides 80-102 and 80-104 were mass produced as GST fusion proteins and removed from the GST partner by thrombin cleavage. They were incubated at RT for 1 hr with primary duck hepatocytes at three different concentrations: 10 μg/ml, 100 μg/ml, and 1 mg/ml. DHBV positive duck serum (1 μl) was then added and incubation continued for three additional hrs. Cells were harvested at day 8 post-infection and intracellular DHBV DNA studied by Southern blot analysis. w/o peptide: without peptide.

Fig. 29(A and B) is a pair of blots showing tissue distribution of p120. A) Detection by the pre-S construct. About 0.5 gm of tissue was homogenized in lysis buffer, precleared and incubated with 2 μg GST fusion protein of 80-102 immobilized on Sepharose beads. Bound proteins were separated by SDS-6% PAGE minigel and visualized by Coomassie blue staining. Position of the p120 band is indicated. B) Detection by direct Western blot. 50 μg of protein was separated by SDS-PAGE and transferred to nitrocellulose filter. The blot was incubated with a rabbit polyclonal anti-p120 antibody and positive signal revealed by HRP-conjugated anti-rabbit serum and DAB.

Fig. 30 is a set of p120 peptides (SEQ ID NOs:70-73).

Fig. 31 is the nucleic acid sequence (SEQ ID NO:74) and deduced amino acid sequence (SEQ ID NO:75) of the p120 and cDNA alone.

Deposit

Under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure, a deposit of the plasmid Ep170pUC has been made with the American Type Culture Collection (ATCC) of Rockville, MD, USA, where the deposit was given Accession No. 69869.

Applicants' assignee, the General Hospital Corporation, represents that the ATCC is a depository affording permanence of the deposit and ready accessibility thereto by the public if a patent is granted. All restrictions on the availability to the public of the material so deposited will be irrevocably removed upon the granting of a patent. The material will be available during the pendency of the patent application to one determined by the Commissioner to be entitled thereto under 37 CFR 1.14 and 35 U.S.C. § 122. The deposited material will be maintained with all the care necessary to keep it viable and uncontaminated for a period of at least five years after the most recent request for the furnishing of a sample of the deposited material, and in any case, for a period of at least thirty (30) years after the date of deposit or for the enforceable life of the patent, whichever period is longer. Applicants' assignee acknowledges its duty to replace the deposit should the depository be unable to furnish a sample when requested due to the condition of the deposit.

Detailed Description

Applicants have identified cellular receptors that specifically interact with the pre-S domain of the DHBV envelope protein. One receptor is a 170 kD cell surface glycoprotein, and is referred to as the p170 receptor. Applicants have mapped the p170 binding site to a major neutralizing epitope of the pre-S domain (amino acids 87-102), within which are two basic amino acids required for virion-receptor interaction. A 46 amino acid pre-S polypeptide covering this binding site inhibits DHBV infection of primary duck hepatocytes. We have also identified an additional pre-S binding protein of 120 kDa (p120). The possible role of p120 as part of the DHBV receptor complex is suggested by its restricted expression in DHBV infectable tissues, by co-localization of its binding site with three virus neutralizing epitopes, and by markedly decreased infectivity of DHBV mutants constructed with impaired p120 binding motif.

Materials and Methods

Cloning and expression of p170 sequences in E. coli. As a negative control, the pre-S domain of DHBV (Fig. 16; Tong et al., Virology 176:596-603, 1990) was amplified by the polymerase chain reaction (PCR) (Saike et al., Science 239:487, 1988), using the sense primer 5'-GCAGATCTATGGGCAGAATCTTTCCAC-3' (SEQ ID NO:1; underlined untemplated BglII site for cloning) and the antisense primer 5'-GTGAATTCAGCGCAGGGTCCCCAAT-3' (SEQ ID NO:2; underlined EcoRI site). Twenty cycles of amplification were carried out using 1 ng of HBV DNA and 1 unit of vent DNA polymerase (New England Biolabs). The PCR product was purified from an agarose gel, digested with BglII and EcoRI, and cloned between the BamHI and EcoRI sites of the pGEX 2TK vector (Pharmacia), which carries coding sequences for the GST protein. DNA fragments of DHBV covering the entire envelope gene (pre-S/S), the pre-S domain, and portions of the pre-S domain were generated by PCR (Fig. 16; Mandart et al., supra). Extra nucleotides for the BglII or BamHI site were put at the 5' end of the sense primers, while extra nucleotides for an in-frame stop codon and an EcoRI site were added at the 5' end of the antisense primer. Twenty cycles of amplification were carried out with one unit of vent DNA polymerase (New England Biolabs) and 1-10 ng of DHBV-16 DNA (Mandart et al., supra). The PCR products were cloned into the BamHI/EcoRI sites of pGEX 2TK. The sense primer for the pre-S region was:

5'-TCAGATCTATGATGGGGCAACATCCAGC-3' (SEQ ID NO:3; underlined BglII site). The antisense primer for the end of the pre-S region: 5'-GCGAATTCAGGTA CCAGACATTTTCTTCTT-3' (SEQ ID NO:4; underlined EcoRI site). The antisense primer for the end of the S region: 5'-GCGAATTCTTATTCCTAACTCTTGTAA-3' (SEQ ID NO:5; underlined EcoRI).

Pre-S deletion mutants are designated by the positions of the first and last pre-S amino acid residues included in the fragment. For example, pre-S(25-104) expressed amino acid residues 25 through 104 of the DHBV pre-S domain. Several pre-S deletion mutants with the 3' ends of the inserts located at the XhoI, SmaI, or HindIII site of DHBV (corresponding to pre-S amino acid 138, amino acid 126, and amino acid 97, respectively) were constructed through double enzymatic digestion of recombinant 2TK plasmids. For example, pre-S(25-126) was generated by removing a short SmaI-EcoRI fragment from pre-S(25-161), with subsequent filling-in of the sticky ends and recircularization of the plasmid by blunt-end ligation. Since a termination codon did not follow the insert immediately, fusion proteins expressed from these particular constructs contained a few miscellaneous amino acid residues at the carboxyterminus derived from the vector sequence downstream of the EcoRI site.

To construct DHBV pre-S substitution mutants, a 1.4 kb EcoRI-BamHI fragment covering the entire pre-S domain was cloned into pAlter-II vector (Promega). Mutagenesis was performed according to supplier's protocol ("Altered sites II: in vitro mutagenesis system technical manual", Promega) from single stranded template DNA. Most mutations introduced or destroyed a restriction enzyme recognition site, so that loss of the restriction site could be monitored as an indication that mutagenesis was successful. Mutations were confirmed by DNA sequencing. The entire pre-S domain of the mutants was amplified and subcloned into pGEX 2TK. Substitution mutants are identified by a single letter for the wild-type amino acid, followed by the numerical position of the amino acid, and a letter for the mutant amino acid. Expression and purification of GST fusion proteins were based on the supplier's protocol (GST Gene Fusion System, Promega). Fusion proteins were expressed by induction with 0.1 mM IPTG for 1 hour and following sonication, purified through a glutathione-sepharose beads (at a ratio of about 1 μl bead per ml bacteria culture). For the expression of the entire

pre-S/S fusion protein of DHBV, induction with IPTG lasted for 3 to 4 hrs. The size, purity and yield of recombinant proteins were analyzed by SDS 12% polyacrylamide gel electrophoresis (PAGE) followed by Coomassie blue staining.

Preparation and labeling of primary duck hepatocytes. DHBV-free Pekin ducklings less than two weeks of age were perfused sequentially with 0.5 mM EGTA and 0.5 mg/ml collagenase through the portal vein (Pugh et al., supra; Tuttleman et al., supra). Hepatocytes were seeded in petri dishes at approximately 90% confluency using Leibovitz's L-15 medium supplemented with 5% fetal calf serum. Subsequent cultures employed serum-free medium supplemented with 1-1.5% dimethyl sulfoxide (DMSO)(Pugh et al., supra). For metabolic labeling, cells were starved in methionine-free Dulbecco's modified Eagle medium (DMEM) for 1 hour, then incubated for 4 hours with L-15 medium supplemented with ^{35}S-methionine (Amersham) or Tran^{35}S Label (New England Nuclear) at 0.1 mci/ml concentration. Cells from each 60 mm dish were treated with 2 ml lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Triton X100, 1% sodium deoxycholate) supplemented with the protease inhibitors aprotinin, leupeptin and PMSF. Nuclei were removed by centrifugation and lysates stored at -80°C. For cell surface labeling with ^{125}I, hepatocytes were detached from a 100 mm petri dish by treatment with Versine/EDTA, washed and resuspended in 0.5 ml phosphate buffered saline solution (PBS). Cell viability was examined by trypan blue exclusion and found to be greater than 90%. Labeling was accomplished with 1 mci of ^{125}I and 40 u of lactoperoxidase (Boehringer-Mannheim), and the labeling reaction was carried out for 20 minutes with four additions of 15 µl 0.04% H_2O_2 (Marchalonis et al. Biochem. J. 113:299-305, 1969). After removing free ^{125}I by centrifugation, cells were lysed as described above.

As a control for cell surface labeling, ^{125}I labeling of unfractionated liver proteins was performed. Frozen duck liver tissue was homogenized in lysis buffer and dialyzed overnight in 1.8% NaCl solution. Around 100 µg protein was labeled with 1 mci of ^{125}I using the iodogen method (Fraker et al., Biochem. Biophys. Res. Comm. 80:849-57, 1978) Labeled proteins were recovered by chromatography through a Sephadex G50 column (Pharmacia).

Detection of pre-S binding proteins in labeled lysates. To reduce the levels of cellular proteins which bind to the GST component of the pre-S fusion protein or directly to the sepharose beads, labeled lysates were first preincubated twice at 4°C with a mixture of empty sepharose beads and GST-bound beads. The beads were washed extensively and used as a negative control in 8% SDS-PAGE. The precleared lysates were then incubated at 4°C for 6 to 16 hr with the specific GST-pre-S fusion protein. After extensive washing of the beads four times with lysis buffer, bound proteins were eluted from beads by heating to 95°C for 5 min, and separated on 8% SDS-PAGE under denaturing conditions. Proteins retained in the second preclearing reaction were run in parallel. The gel was fixed with 10% acetic acid, treated with Amplify solution (Amersham), dried, and exposed. For experiments performed with the ^{125}I labeled proteins, the treatment step with Amplify solution was omitted. Comparison of protein bands binding only to the pre-S fusion indicated the specificity of the interaction.

Infection of primary duck hepatocytes and polypeptide inhibition assay. Pre-S polypeptides expressed as GST fusion proteins were purified on glutathione sepharose beads and cleaved with thrombin. The beads which contained the GST moiety were removed and supernatant collected. Primary duck hepatocytes were preincubated with different concentrations of proteins at room temperature for 30 minutes.

DHBV positive duck serum ('viremic sera') was obtained from a Pekin duckling transfected with cloned DHBV DNA. Two weeks after transfection, when DHBV became strongly positive in the blood, the duckling was bled and blood stored at 4°C overnight. The blood was centrifuged to obtain sera. Five µl of the resulting DHBV positive duck serum were added and incubation continued at 37°C for three hours. The cells were extensively washed and incubated with fresh L-15 medium supplemented with a 1:200 dilution of rabbit anti-pre-S antiserum to prevent secondary infection. This antiserum was obtained by immunizing a rabbit with purified pre-S polypeptide (amino acids 1-161, expressed as GST fusion protein and cleaved to remove GST). Cells were harvested one-week postinfection and assayed for intracellular DHBV nucleic acid by Southern blot analysis.

Detection of p170 in different duck tissues. For these experiments, the elution profile of p170 through an anion exchange column was established. One ml of 35S-

labeled duck hepatocyte lysates was dialyzed overnight in 50 mM Tris-HCl, pH 8.3. The lysates were applied to a column packed with pre-swollen DEAE-cellulose (Sigma) equilibrated with 50 mM Tris-HCl, pH 8.3. Bound proteins were sequentially eluted with 100, 200 and 400 mM NaCl in 50 mM Tris-HCl, pH 8.3. The peak of radioactivity in each fraction was collected and dialyzed back against lysis buffer. After incubation with GST-pre-S fusion protein, bound proteins were revealed by 8% SDS-PAGE and fluorography. To study tissue distribution of p170, 0.6-1.2 g of frozen tissue was homogenized in 6-12 ml of lysis buffer and after overnight dialysis against 50 mM Tris-HCl, pH 8.3, insoluble materials were removed by centrifugation (10,000 g for 20 min. at 4°C) followed by filtration through a 0.45 μm filter. The solution was passed through a column containing 8 g of preswollen DEAE-cellulose and eluted with 100 mM and 200 mM NaCl in 50 mM Tris-HCl, pH 8.3. The 200 mM NaCl eluent was dialyzed against the lysis buffer and precleared twice with 10-20 μl bed volume of Sepharose beads. The concentration of proteins in each sample was determined by the Biorad protein assay (Lowry et al., J. Biol. Chem. 193:265, 1951), and 7 mg protein from each sample was incubated with 4-8 μg of GST-pre-S fusion protein. After separation of bound proteins with a 8% SDS-PAGE, protein bands were visualized by silver staining, using a Gelgold staining kit (Pierce).

Protein microsequencing of p170. p170 was purified from 40 g of duck liver using the method described above, separated from GST-pre-S protein by SDS-PAGE, and transferred to polyvinylidene difluoride membrane (PVDF) membranes (Biorad). After staining with Ponceau S, strips of the membrane containing about 20 μg of p170 were obtained for sequencing analysis by digesting with lyase C, and separating high pressure liquid chromatography (HPLC). Selected peptide peaks were sequenced by the Edman degradation method.

Cloning of p170 cDNA. An intra-peptide "miniPCR" procedure was used to obtain portions of the coding sequences, designated peptide 1 and peptide 3. The degenerate nucleic acid sequences encoding peptides 1 and 3 were then used as unique PCR primers to amplify the coding region spanning these two peptides. The degenerate PCR primers were:

a) peptide 1 sense: 5'-GAXYTNTAYGTNATGGAGAT-3' (SEQ ID NO:6)

b) peptide 1 antisense: 5'-AAYTCNGGYTCNCCNGCYTCXTG-3' (SEQ ID NO:7)

c) peptide 3 sense: 5'-TZYTNAGYCAYGAXTTYCAXG-3' (SEQ ID NO:8)

d) peptide 3 antisense: 5'-TTZGCNGAXTANAXNGTYTC-3' (SEQ ID NO:9)

Fig. 15 shows amino acid residues 6-27 of peptide 1 (pk-81), and the entire 22 amino acids of peptide 3 (pk-69). Potential nucleotide sequences that code for peptides 1 and 3 are shown below each amino acid sequence. "Degenerate primer & product" identifies the primers used for intra-peptide miniPCR. "Specific primer & product" refers to specific amplification of 2.5 kb sequences between peptide 1 and 2 using a sense primer derived from peptide 1 and an antisense primer derived from peptide 3. Primers are shown in small letters, while amplified sequences are shown in capital letters. For convenience the antisense primers are written in the sense orientation.

The template was first strand cDNA transcribed from duck liver mRNA using random hexamer primers and superscript II reverse transcriptase (Gibco/BRL). Thirty-five cycles of amplification consisting of denaturing and annealing steps were carried out. The PCR products of expected sizes were isolated from a 4% NieSieve agarose gel, cloned, and sequenced.

Results

A 170 kd cell surface glycoprotein binds to the pre-S domain of DHBV large envelope protein. The entire envelope protein of DHBV and the pre-S domain of HBV (the combined pre-S1 and pre-S2 domains) were cloned into the pGEX 2TK vector and expressed as GST fusion proteins. Fig. 1A shows the expression and purification of the GST fusion proteins in which GST is fused to: the DHBV pre-S/S protein (lane 1); the DHBV pre-S domain (lane 2); the HBV pre-S domain (lane 3); or the intact GST protein expressed from the pGEX 2TK expression vector (lane 4). In addition to a protein band of expected size, all the recombinant constructs displayed a band with a mobility slightly faster than the intact GST protein. The band was seen in all the additional fusion protein constructs (Fig. 7A and 6B). It most likely corresponds to a proteolytic cleavage product of the fusion proteins around the cloning site.

Using the DHBV pre-S/S protein fused to GST as a probe, a 170 kd glycoprotein was identified in ^{35}S methionine labeled duck hepatocyte lysates that interacted with the DHBV pre-S domain (Fig. 1B, lane 3). In Fig. 1B, ^{35}S labeled primary duck hepatocyte

lysates were precleared with Sepharose beads and GST-bound beads (lane 1), and then incubated with different GST fusion proteins immobilized on Sepharose beads. Fig. 1B shows the proteins isolated by binding to the following GST fusion protein: GST-DHBV pre-S/S protein (lane 3); GST-DHBV pre-S domain (lane 4), and GST-HBV pre-S domain (lane 2). The bound liver proteins were separated on SDS-PAGE and identified by fluorography. In a separate experiment, hepatocytes were metabolically labeled in the presence of tunicamycin. After preclearing (lane 7), the lysates were reacted with GST fusion proteins of the DHBV pre-S domain (lane 5) or the HBV pre-S domain (lane 6). Since the p170 protein was not retained by GST alone (lane 1), it appeared specific for DHBV sequences. The p170 protein was also retained by GST-pre-S fusion protein (lane 4), suggesting that the S domain was not essential for binding activity. Failure of a similar pre-S fusion protein derived from HBV to bind duck hepatocyte p170 (lane 2) as a control is consistent with a species specificity of the p170-pre-S interaction.

To determine whether p170 is glycosylated, duck hepatocytes were labeled in the presence of tunicamycin (2 μg/ml). Under these conditions, a protein of approximately 145 kd was detected (lane 5).

The p170 receptor is located on the cell surface of hepatocytes. To show this, cell surface proteins of primary duck hepatocytes were labeled with ^{125}I using the lactoperoxidase reaction technique. The ^{125}I surface labeled lysates were precleared (lane 6) and incubated with GST fusion proteins of the whole DHBV pre-S domain (lane 1) or of various deletion mutants of the DHBV pre-S domain fused to GST: GST-pre-S(25-102) (lane 2), GST-pre-S(80-104) (lane 3), GST-pre-S(80-102) (lane 4), or GST-pre-S(92-161) (lane 5).

A control experiment was performed to show the specificity of cell surface labeling by lactoperoxidase. In Fig. 2, lanes 1 and 2 show cell surface proteins labeled with ^{125}I by lactoperoxidase. Lanes 3 and 4 show lysates of liver tissue labeled with ^{125}I by iodogen. After a brief preclearing, ^{125}I labeled lysates were incubated with GST fused DHBV pre-S domain (lanes 1 and 3) or a deletion mutant D80-102, which did not bind p170 (lanes 2 and 4). Note that proteins that were nonspecifically bound to Sepharose beads were different using the two labeling methods. Compared to direct ^{125}I labeling of unfractionated liver tissues by the iodogen method, cell surface labeling with

lactoperoxidase produced a different pattern of nonspecific binding proteins (Fig. 2, compare lanes 1, 2 with 3, 4). A 170 kd molecule reacting specifically with GST-pre-S protein was nevertheless detected under both labeling conditions (Fig. 2, lanes 1 and 3; Fig. 1C, lane 1).

Reactivity of this molecule with several pre-S deletion mutants confirmed it to be the same p170 as detected in metabolically labeled lysates. Mutants pre-S(25-102) and pre-S(80-104), which bound to and retained p170 from ^{35}S labeled lysates (see below), also retained the 170 kd protein from surface labeled lysates (Fig. 1C, lanes 2 and 3). On the other hand, mutants 80-102 and 92-161, which failed to bind p170 in ^{35}S labeled lysates, were also unable to bind the 170 kd cell surface protein (lanes 4 and 5). Thus, p170 appears to be present on the cell surface of hepatocytes.

Interaction of DHBV particles with p170 in vitro and in vivo. Since it was possible that the fusion proteins expressed in E. coli may not have had the same conformation as the corresponding native viral particles, it was important to determine whether p170 could bind to native viral particles. A competition experiment was performed between native DHBV particles and GST-pre-S/S fusion protein by adding highly viremic duck serum into the incubation reaction (Fig. 3A, Fig. 3B). Three ml of ^{35}S labeled lysates were precleared (lane 8). An equal volume of each lysate was incubated with 5 μg of GST-DHBV pre-S/S fusion protein in the absence of duck sera (lane 1), or in the presence of 60 μl (lane 2) or 200 μl (lane 3) of DHBV-free duck sera, or in the presence of 5 μl (lane 4), 20 μl (lane 5), 60 μl (lane 6), or 200 μl (lane 7) of DHBV viremic sera. Bound proteins were revealed by SDS-PAGE and fluorography. Incubation of ^{35}S labeled hepatocyte lysates with 5 μg of GST-pre-S/S fusion protein gave rise to a strong band of p170 (Fig. 3A, lane 1). Addition of 60 μl (lane 2) or 200 μl (lane 3) DHBV-free duck serum to the incubation mixture had little effect on binding to p170. In contrast, as little as 5 μl of DHBV positive serum strongly inhibited p170 binding (lane 4). Increasing the volumes of the DHBV positive sera diminished the p170 band in a dose-dependent manner (lanes 5, 6, and 7). With an incubation of 200 μl of viremic serum, virtually no p170 binding was observed (lane 7). A similar inhibitory effect was shown when DHBV particles purified through successive sucrose gradient

centrifugation were applied. These results demonstrate that p170 is recognized by virion particles through the pre-S domain.

If DHBV particles bind p170 in vitro, they might also do so in vivo during natural "wild-type" viral infection. Three 100 mm dishes of primary duck hepatocytes were prepared from a DHBV-free duckling. One dish served as a control while the other two dishes were infected overnight with 30 µl and 300 µl of DHBV viremic serum, respectively. Nine days post-infections, cells were metabolically labeled and lysed. In Fig. 3B, precleared lysates were incubated with GST-pre-S fusion protein. Lane 1: noninfected cells; lane 2: cells infected with 30 µl viremic serum; lane 3: cells infected with 300 µl viremic serum. Lanes 4 and 5: comparison of p170 expression in duck hepatocytes (lane 4) and fibroblast-like cells derived from duck hepatocytes (lane 5). Hepatocytes were cultured in L-15 medium supplemented with 1% DMSO (lane 4) or 5% fetal calf serum (lane 5) for 10 days before labeling with 35S methionine. Lysates, prepared from these hepatocytes were reacted with GST-pre-S fusion protein.

Although p170 was detected as a strong band in uninfected hepatocytes (Fig. 3B, lane 1), it was not detected in the two infected dishes (lanes 2, 3). Similar results were obtained with 125I-labeled cell surface protein lysates. Thus, either newly synthesized p170 was masked by binding to endogenous DHBV envelope protein or de novo synthesis of p170 was severely inhibited during infection.

To compare the steady-state levels of p170 between infected and noninfected duck livers, unlabeled liver tissues from two-week-old ducklings were studied. One duckling was DHBV-free, while the other was naturally infected. Liver tissues were homogenized in lysis buffer and precleared twice with empty Sepharose beads. After incubation with GST-pre-S fusion protein immobilized to Sepharose beads, bound proteins were visualized by SDS-PAGE and Coomassie blue staining of the gel. The results are shown in Fig. 4 (lane 1: 2 g of DHBV-free liver; lane 2: 2 g of infected liver; lane 3: 2 g of DHBV-free liver mixed with 1 ml of DHBV + serum from the infected duck; lane 4: 1 g of DHBV-free liver mixed with 1 g of infected liver). When the same amount of tissue was used, p170 was readily detected in uninfected liver but barely visible in infected liver (lanes 1 and 2). Adding either DHBV-positive duck serum or lysates of the infected liver masked p170 from DHBV-free duckling (lanes 3 and 4).

Thus, p170, even if present at normal levels, would be rendered undetectable by a large number of virus particles.

Tissue distribution of p170. To examine whether expression of p170 is dependent on the differentiated status of duck hepatocytes, hepatocytes were cultured either in serum-free, DMSO-containing L-15 medium as described, or in L-15 medium supplemented with 5% calf serum instead of DMSO. Incubation in the L-15 medium causes rapid loss of both hepatocyte morphology and susceptibility to DHBV infection (Pugh et al., supra). After ten days of culture, hepatocytes maintained with calf serum became largely elongated so as to resemble fibroblasts. However, when cells were metabolically labeled and then cell lysates were incubated with pre-S fusion protein, the intensity of the p170 band was virtually unaffected (Fig. 3B, compare lanes 4 and 5). Thus, expression of p170 does not depend on differentiation of the hepatocyte.

To study further the tissue specificity of p170 expression, a two-step purification procedure involving anion exchange column and affinity chromatography was developed. ^{35}S labeled hepatocyte lysates were run through a DEAE-cellulose column. The flow through fraction and eluent fractions (100 or 200 mM NaCl) (200 mM NaCl) were incubated with GST-pre-S fusion protein. Bound proteins were revealed by SDS-PAGE and fluorography. The results are shown in Fig. 5A (lane 1: flow through fraction; lane 2: 100 mM NaCl eluent; lane 3: 200 mM NaCl eluent). The position of p170 is shown by an arrow. These experiments showed that p170 was eluted with 200 mM NaCl from a DEAE-cellulose column (Fig. 5A, lane 3). This fractionation procedure removed the majority of nonspecific binding proteins.

To study the tissue distribution of the p170 protein, 0.6-1.2 g of either tissue lysates or serum derived from DHBV-free ducklings were passed through a DEAE-cellulose column. The 200 mM NaCl eluents were precleared with Sepharose beads. Seven mg of protein (7 mg for lysates and 70 mg for serum) were incubated with immobilized pre-S fusion protein. Bound proteins were separated by 8% SDS-PAGE and detected with a Gelgold silver staining kit (Fig. 5B). The lanes of Fig. 5B are: lane 1: 2 μg of the 170-kd molecular size marker α2-macroglobulin; lane 2: GST-pre-S fusion protein (the same amount as used in purifying p170 from each tissue); lanes 3-13: purification and detection of p170 in heart (3), lung (4), liver (5), muscle (6), spleen (7),

stomach (8), gall bladder (9), kidney (10), 7 mg serum protein (11), 70 mg serum protein (12), pancreas (13). Lanes 14 and 15 contain the second preclearing reaction for stomach (14) and gall bladder (15). Lanes 13-15 were derived from a separate SDS-PAGE and staining.

SDS-PAGE and silver staining revealed that p170 was highly expressed in pancreas, liver, kidney, and spleen (Fig. 5B, lanes 13, 5, 10, and 7, respectively). These are the same tissues in which DHBV replication has been reported. The p170 receptor was also found in lung, heart, and, to a less extent, in stomach and muscle tissue (lanes 4, 3, 8, and 6, respectively). In gall bladder, the major binding protein for the pre-S protein had a molecular size of around 180 kd, and was very abundant (lane 9). This 180 kd band was not seen in the preclearing reaction (lane 15). The p170 receptor was not detected in serum (lane 11), even when a 10-fold excess concentration of proteins was applied to the gel lane (lane 12). Thus, p170 does not appear to be a secreted protein.

p170 binds to a major neutralizing epitope of the DHBV pre-S domain. In order to define the region of the pre-S protein that is essential for binding to p170, nine progressive aminoterminal and five carboxyterminal deletion mutants were made and expressed as GST fusion proteins (Fig. 6). Fig. 6 shows the 161 amino acid residues of DHBV pre-S domain. Positions are given for the first amino acid residues in the aminoterminal deletion mutants, the last residues in the carboxyterminal deletion mutants, and both terminal residues in the double deletion mutants. Positive (+) or negative (-) binding results with p170 are shown to the right of each mutant. For both Fig. 7A and Fig. 7B, lanes 1 through 21 correspond to the intact pre-S domain (1), or to the deletion mutants 25-161 (2), 59-161 (3), 71-161 (4), 80-161 (5), 84-161 (6), 87-161 (7), 92-161 (8), 98-161 (9), 112-161 (10), 1-98 (11), 1-102 (12), 1-104 (13), 1-126 (14), 1-138 (15), intact pre-S (16), 25-104 (17), 25-102 (18), 59-102 (19), 80-102 (20), and 80-104 (21). Lane 22 in Fig. 4B shows the intact GST protein.

All of the fourteen deletion mutants expressed GST fusion proteins of the expected sizes as judged by SDS-PAGE (Fig. 7A). Removal of up to 86 amino acid residues in the aminoterminus reduced, but did not abolish, binding (mutant 87-161; Fig. 7B, lane 7). Deleting five additional amino acid residues abolished binding (92-161; lane 8). A mutant with a carboxyterminal deletion of up to 59 residues retained strong binding

capacity (1-102; lane 12), while further deletion to amino acid 98 abolished binding (1-98; lane 11). Therefore, the pre-S sequence critical for p170 binding was localized to the 16 amino acid sequence between residues 87 and 102.

This sequence covers a known virus-neutralizing epitope (type II) that maps to amino acids 91-99 (Yuasa et al., Virology 181:14-21, 1991) and overlaps with two additional neutralizing epitopes located at amino acids 83-90, and 100-107, respectively (Fig. 8). The locations of type II and IV epitopes indicated in Fig. 8 are according to Yuasa et al, supra; the location M-900 and SD20 epitopes are according to Chassot et al., supra. The location of a sequence nonessential for viral infectivity is according to Li et al., supra.

To test whether the minimum binding sequence functions independently of other pre-S amino acid regions, five double deletion mutants were constructed (Fig. 6 and Fig. 7A). Of the three double deletion mutants terminating at amino acid 102, only the one with a limited (24 amino acid) aminoterminal deletion bound p170 (25-102; Fig. 7B, lane 18). On the other hand, both mutants terminating at amino acid 104 were able to bind p170 (lanes 17 and 21). The shortest double deletion mutant capable of binding p170 was the construct 80-104, which contained only a 25 amino acid sequence (lane 21).

To further test the role of the pre-S amino acid sequence around the type IV epitope in the p170 interaction, five additional double deletion mutants were constructed: pre-S(25-126), pre-S(59-126), pre-S(71-126), pre-S(42-102), and pre-S(59-104) (Fig. 9). Fig. 6B is an autoradiograph showing the retention of p170 by different mutants (lane 1: preclearing reaction; Lanes 2-5: mutants terminating at residue 126; lane 2: 1-126; lane 3: 25-126; lane 4: 59-126; lane 5: 71-126; Lanes 6-8: mutants terminating at residue 102: lane 6: 25-102; lane 7: 42-102; lane 8: 59-102. Lanes 9-11: mutants terminating at residue 104; lane 9: 25-104; lane 10: 59-104; and lane 11: 80-104.

Of the four mutants sharing an identical carboxyterminus at residue 126, but differing in their aminotermini (residues 1, 25, 59, and 71, respectively), p170 retention was not reduced, but rather increased, by successive aminoterminal truncations (Fig. 10, lanes 2-5). For the three double deletion mutants terminating at amino acid 104, moving the amino terminal end from amino acids 25 to 59 to 80 likewise did not reduce p170 binding (lanes 9, 10, and 11). Only for the mutants with a carboxyterminus at amino acid

102 did a change in the N-terminus from amino acid 25 to amino acid 42 greatly reduce p170 binding (lanes 6 and 7). Moving the aminoterminus further down to amino acid 59 completely abolished binding (lane 8), as shown in Figs. 6, 7A, and 7B.

The p170 binding site in the human HBV pre-S protein. Fig. 20 shows the correspondence between the p170 binding site in the DHBV pre-S protein and a predicted p170 binding site in the human HBV pre-S protein. When pre-S residues 71-118 of HBV were aligned against residues 69-116 of DHBV, there were numerous identical residues (shown by asterisks) and no gaps in the sequence alignment. Thus, the p170 binding site at residues 87-102 of DHBV corresponds to a predicted p170 binding site at residues 89-104 of HBV. The arginine residue at position 97 of DHBV, which is critical for p170 binding, is conserved in HBV at HBV pre-S residue 99.

A pre-S polypeptide covering the p170 binding site inhibits DHBV infectivity. The coincidence of the location of the p170 binding site with the major pre-S neutralizing epitopes raises the possibility that p170 is the primary DHBV receptor required for viral entry into hepatocytes. However, the affect of neutralizing antibodies could be caused by steric hindrance or conformational change instead of by direct attachment to the receptor binding site. To directly test the importance of p170 binding by the pre-S domain on viral infectivity, an infection inhibition experiment was performed. The pre-S polypeptides were expressed as GST fusion proteins as described above, and the GST domain was removed by thrombin cleavage. Primary duck hepatocytes were preincubated at room temperature for 30 minutes with three different concentrations of pre-S polypeptides pre-S(59-104) or pre-S(59-102), which when expressed as GST fusion proteins are capable of binding the p170 receptor. Cells were then infected at 37°C for 3 hours with 2 µl of viremic duck serum. Unattached virions were washed away and any secondary round of viral infection was prevented by the constant presence of rabbit anti-pre-S neutralizing serum. Viral DNA in hepatocytes was analyzed by Southern blot one week post-infection. The results are shown in Fig. 11 (Lane 1: Control hepatocytes infected in the absence of pre-S polypeptide; lanes 2-4: peptide 59-104 at 10 µg/ml (1), 100 µg/ml (2), and 1 mg/ml (3)). Peptide pre-S(59-104), whose GST-fusion protein binds p170, inhibited DHBV infection at 100 µg/ml and 1 mg/ml concentrations. These results demonstrate that the p170 binding site is a viral receptor binding site.

Effect of single amino acid changes in the major pre-S neutralizing epitope region on p170 binding. In order to define the individual amino acid residues critical for p170 binding, site-directed mutagenesis experiments were carried out on amino acids 88 through 102 of the pre-S domain. Eleven single amino acid substitution mutants involving nine amino acid residues were constructed. Fig. 12 shows the nature and location of amino acid changes in these mutants. Amino acid sequences from position 87 through 102 are shown for wild-type DHBV. Underlines denote amino acid residues that are variable in a goose hepatitis B virus strain (Shi et al., GenBank accession number M95589). Bold-face letters denote residues conserved in a comparison with heron hepatitis B virus (Sprengel et al., J. Virol. 62:3832-3839, 1988).

The affinity of the pre-S mutants for p170 binding was compared with that of the "wild-type" fusion protein (Fig. 12 and Figs. 13A and 13B). Fig. 13A shows the level of expression of GST-pre-S fusion proteins from the mutants. Proteins purified from equal amounts of bacterial culture were applied to 12% SDS-PAGE. Mutants were arranged in the following order: E91G (lane 1), E92V (2), D93F (3), K95S (4), R97L (5), R97C (6), E98A (7), E98V (8), W88S (9), P9OL (10), R102G (11), K95S/R97C (12), K95S/R97L/E98A (13), K95S/A96T (14), E91G/K95S (15), E91G/R97C (16). Protein size markers are shown to the right of the figure. The levels of pre-S fusion proteins produced by the mutant constructs were similar (Fig. 13A) and equal amounts of the fusion proteins were used for binding experiments.

Fig. 13B shows the binding capacities of each pre-S substitution mutant for p170. Lanes 1 through 18 are: second preclearing (1), wild-type pre-S fusion protein (2), W88S (3), P9OL (4), E91G (5), E92V (6), D93F (7), K95S (8), R97L (9), R97C (10), E98A (11), E98V (12), R102G (13), E91G/K95S (14), E91G/R97C (15), K95S/A96T (16), K95S/Rp7L/E98A (18) While all of the mutants exhibited reduced retention of p170, mutants R97L and R97C consistently showed the lowest binding activity in several independent experiments. Mutant K95S also had greatly reduced binding capacity (lane 8). These results show that these two basic amino acid residues are important in the interaction of the pre-S domain with p170. However, when different amino acid substitutions were combined to produce double or triple amino acid changes (Fig. 12 and

Fig. 13A), none showed a further decrease in p170 binding, even for double mutations at both residues 95 and 97 (K95S/R97C and K95S/R97L/E98A; Fig. 13B, lanes 17, 18).

p170 is structurally related to carboxypeptidases. To gain insight into the molecular identity of p170 and to provide essential peptide sequence for cloning the p170 cDNA, p170 receptor was purified from DHBV-free duck liver using a combination of ion-exchange chromatography and affinity chromatography. Purified proteins were separated by SDS-PAGE and transferred to PVDF membranes. The 170 kd protein band was digested with lyase C, and selected peptides were sequenced. Four examples of p170 peptide sequences are shown in Fig. 14A. A search of a computerized data base for peptides 3 and 4 did not reveal any significant degree of similarity to any known protein. However, peptides 1 and 2 are similar to mammalian carboxypeptidases H, N, and M, the strongest binding being to carboxypeptidase H. Fig. 14B shows the similarity of peptide 1 to basic carboxypeptidases. Fig. 14C shows the similarity of peptide 2 to basic carboxypeptidases. The dots denote residues identical to the peptide sequence. The acronyms shown in Figs. 14B and 14C are: CPH (carboxypeptidase H); CPM (carboxypeptidase M); CPN (carboxypeptidase N); and AEBP1 (a mouse transcriptional repressor with carboxypeptidase activity). The nucleotide and amino acid sequence of each carboxypeptidase is known: bovine CPH (Fricker et al., Nature, 323:461-64, 1986); human CPH (Manser et al., 267:517-525, 1990); human CPM (Tan et al., M. J. Biol. Chem. 264:13165-13170, 1989); human CPN (Gebhard et al., Eur. J. Biochem., 178:603-07); and AEBP1 gene (GenBank accession number X80478).

Cloning of p170 cDNA. Degenerate PCR primers were synthesized to amplify the middle portion of coding sequences for peptides 1 and 3, respectively (see Materials and Methods).

Sequencing of the PCR products unveiled the coding sequences for residues 12-18 of peptide 1 and residues 9-15 of peptide 3 (ATGAAACAGACACTGAAGAA; SEQ ID NO:10) (Fig. 15). This information enabled synthesis of unique PCR primers to amplify the region between these two peptides. With a sense primer derived from peptide 1 (ATGGAGATCTCGGACGGCCC; SEQ ID NO:11) and an antisense primer from peptide 3 (TTCTTCAGTGTCTGTTTCAT; SEQ ID NO:12), a 2.5 kb cDNA was amplified from randomly primed duck liver cDNAs. Sequencing the 5' end of PCR

clones revealed nucleotides coding for residues 16-27 of peptide 1. Sequencing the 3' end identified coding sequence for residues 1-9 of peptide 3 (Fig. 15). These results are compatible with the 2.5 kb cDNA being part of the p170 cDNA.

Figs. 18 and 19 show the partial nucleotide and deduced amino acid sequences of the 2.5 kb p170 cDNA. The sequence of Fig. 18 corresponds to 1.1 kb of nucleotide sequence at the 5' end of the 2.5 kb p170 cDNA. The peptide 1 amino acid sequence corresponds to amino acid residues 1-12 of the sequence of Fig. 18. The peptide 4 amino acid sequence corresponds to residues 337-349 of the sequence of Fig. 18. The presence of the peptide 4 coding sequence in the 2.5 p170 cDNA confirms it as the authentic cDNA for the p170 receptor. The sequence of Fig. 19 corresponds to 460 bp at the 3' end of the 2.5 kb p170 cDNA. The sequence of peptide 3 corresponds to the carboxyterminal 15 amino acid residues of the sequence of Fig. 19.

The 2.5 kb cDNA was blunt-ended and cloned into plasmid pUC18 cut with SmaI. The plasmid was transformed into E.coli strain DH5α. After overnight growth on LB plates supplemented with ampicillin, IPTG, and X-gal, several white colonies were isolated and grown in liquid LB medium supplemented with ampicillin. A positive clone harboring a plasmid containing a 2.5 kb insert was named Ep170pUC and deposited with the ATCC with designation No. 69869.

Pre-S fusion protein constructs leading to identification of p120. As described above, to facilitate the identification of binding proteins for the DHBV pre-S molecule and mapping of the binding site, full-length and truncated forms of pre-S protein were expressed as fusion constructs with glutathione S transferase (GST) and immobilized on glutathione-Sepharose beads. Radiolabeled liver proteins which bound to the pre-S protein were retained on the beads and could be subsequently visualized by SDS-PAGE and fluorography. Some of the truncated pre-S constructs used to identify p120 are described in detail above in connection with p170. Additional deletion mutants were generated by twenty cycles of PCR amplification of a DHBV clone 16 and inserted into the BamHI-EcoRI sites of pGEX 2TK vector (Pharmacia). A stop codon was incorporated into each antisense primer to ensure a pre-determined C-terminus. Deletion mutants were named by positions of the first and last amino acid residues of the pre-S sequences.

To generate mutant E98S-102, plasmid DNA of mutant 25-102 cloned in 2TK was double digested with BamHI (cut at the 5' end of pre-S insert) and HindIII (cut at sequence coding for pre-S residue 98) to remove 0.2 kb coding sequence. After filling the cohesive ends with Klenow fragment and dNTP, the plasmid DNA was recircularized. To introduce single amino acid substitutions into residues 98-102 in the construct 80-102, specific mutagenic antisense primers were used in the PCR. Incorporation of point mutations into construct 92-161 required two steps. Mutation was first introduced into a 1.4-kb EcoRI-BamHI fragment of DHBV genome cloned into the pALTER-II mutagenesis vector (Promega). DNA fragments coding for residues 92-161 were then amplified by 20 cycles of PCR and cloned into pGEX-2TK vector.

Detection of pre-S binding proteins. Preparation of primary duck hepatocytes, metabolic labeling, and detection of pre-S binding proteins leading to identification of p120 were carried out as described above in connection with p170. To examine the availability of p120 on cell surface, plated hepatocytes on petri dishes were washed three times with PBS, incubated at room temperature (RT) for 30 min with 1 mM sulfo-LC-biotin (Pierce) in PBS, and washed again three times with PBS before lysis of cells. The precleared lysates were incubated with various constructs of pre-S fusion protein immobilized on Sepharose beads and retained materials were separated on SDS-PAGE. Proteins were transferred to nitrocellulose membrane and nonspecific binding sites blocked at RT for 1 hr with 5% BSA in PBS-0.05% Tween 20 (PBST).

After incubation at RT for 1 hr with a 1:2000 dilution of streptavidin conjugated with horse radish peroxidase (HRP; Pierce), biotinylated proteins were visualized by enhanced chemiluminescence (ECL). As a negative control, immunoprecipitation of a cytoplasmic protein (Golgi β-COP) was performed on both surface labeled hepatocyte lysate and labeled total liver proteins. For biotinylation of total liver proteins, 1 ml of liver tissue lysate (corresponding to 100 mg liver tissue) was incubated with 2.5 mg sulfo-LC-biotin for 1 hr at 40°C. The reaction was terminated by addition of glycine to 100 mM followed by further incubation for 3 hr. 6 mg of surface or total labeled liver proteins was incubated with 1:50 dilution of a monoclonal antibody against Golgi β-COP protein (M3A5, Sigma) for 2 hr at 4°C. After addition of 100 ml protein G-Sepharose

beads (50% slurry), samples were incubated overnight. The retained proteins were analyzed as described above.

To study the tissue distribution of p120 by means of its affinity with truncated pre-S protein, 0.5 gm of frozen tissue was homogenized in 10 ml of lysis buffer, precleared twice with Sepharose-GST beads, and then incubated with 2 mg of immobilized GST fusion protein of pre-S construct 80-102. After separation of bound proteins with a SDS-6% PAGE minigel (BioRad), protein bands were visualized by staining with Coomassie blue. To detect small amount of p120 retained by the intact pre-S protein (1-161), bound proteins separated by SDS-8% PAGE were transferred to nitrocellulose filter. The blot was incubated successively with 1:1000 dilution of a rabbit polyclonal anti-p120 antiserum (16a) in PBST and 1:1000 dilution of a donkey anti-rabbit Ig conjugated with HRP, and positive bands revealed with Sigma Fast DAB (3, 3'-Diaminobenzidine) tablets dissolved in water.

To detect p120 in duck tissues by direct Western blot, 50 mg of protein was separated by 6% SDS-PAGE and transferred to nitrocellulose filter. The filter was blocked with 3% BSA in PBST, and incubated with a 1:1000 dilution of a rabbit polyclonal anti-p120 antibody at RT overnight. After a thorough wash, the filter was incubated with 1:1000 dilution of a donkey anti-rabbit Ig conjugated with HRP, and positive bands revealed with Sigma Fast DAB.

Inhibition of DHBV infectivity by pre-S p120 peptides. Primary duck hepatocytes were seeded at a density of $3x10^5$ cells/well into 12-well plates. The experiments were performed within one week after plating. Pre-S polypeptides 80-102 and 80-104 expressed as GST fusion proteins were purified onto Sepharose beads. The GST portion was removed by digestion with thrombin (Sigma) and centrifugation. Thrombin (MW 55 kDa) was not removed in initial experiments but removed in repeat experiments through centricon 30 filter (cut-off 30 kDa), and similar results were obtained. Because of the nature of gene fusion with the 2TK vector, all the peptides contained at their N-terminus nine irrelevant amino acid residues (GSRRASVGS; SEQ ID NO:13) contributed by thrombin recognition site, protein kinase domain, and BamHI site.

Hepatocytes in the plates were first incubated with the peptides at three different concentrations at RT for 1 hr, followed by addition of 1 μl viremic serum and a further incubation for 3 hrs. After extensive wash, cells were maintained in L15-1% DMSO medium supplemented with a neutralizing rabbit polyclonal anti-pre-S antiserum (16a) to suppress virus spread. Cells were harvested at day 8 post-infection. The experiments were performed in duplicate and harvested cells pooled for hybridization experiments.

Transfection and infection with p120 substitution mutants. pDHBV3.5 was constructed to contain a 3.5-kb overlength DHBV genome (NcoI-NsiI fragment of DHBV clone 16; reference 19) inserted between the EcoRI and PstI sites of pUC18 vector. This construct contained the coding sequence for viral pregenomic RNA and therefore produced infectious DHBV particles when transfected into LMH cells. Mutagenesis was performed on the 1.4-kb EcoRI-BamHI fragment of DHBV as described above and mutations were confirmed by DNA sequencing. The 0.8-kb BglII-XhoI fragment was excised from the mutagenesis vector to replace the corresponding fragment in the wild-type pDHBV3.5. Plasmid DNA used for transfection experiments were isolated from bacterial culture by the Promega Maxi prep columns and further purified by ultracentrifugation through CsCl gradient. DNA (20 mg) was transfected into a 100 mm dish of LMH cells by the calcium phosphate method. The cells were maintained in 1:1 ratio of DMEM/F12 medium supplemented with 10% fetal bovine serum. Each mutant construct was transfected in duplicate and materials pooled from the two dishes were used for hybridization. Cells were harvested at day 7 post-transfection and levels of DHBV cccDNA, total and core RNA, and core DNA were detected by Northern and Southern blot analyses. Secreted core particles and virion particles were concentrated from culture medium by centrifugation through a 10-20% sucrose gradient or by precipitation with polyethyl glycine (PEG 8000, final concentration 10%). After successive digestion with DNase I, proteinase K followed by phenol/chloroform extraction, amounts of DHBV DNA was determined by Southern Blot. To selectively detect virion particles, the pelleted material was digested with Pronase and DNase I.

Virion particles concentrated from to 3 to 6 ml of culture medium were used to infect primary duck hepatocytes cultured in 6-well plates. After incubation at 37°C for

6.5 hr, cells were washed and cultured for additional 7 days before extraction of total cellular DNA for Southern blot analysis.

Purification of p120 from duck liver. A total of forty grams of frozen liver from two-week-old Pekin ducklings were homogenized in 300 ml lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate) using a Polytron homogenizer. The homogenate was clarified by low speed spinning followed by 10 min centrifugation at 14000 rpm in a Sorval SS34 rotor. The supernatant was filtered through 0.45 μm filter units and each 50 ml aliquot was precleared once with 80 μl (10% v/v) killed Staphylococci aureus positive for protein A (Boehringer Mannheim), and once with a mixture of 100 μl bed volume (b.v.) glutathione Sepharose beads and 5 ml b.v Sepharose beads conjugated with GST protein, at 4°C for 8 hr to overnight. Proteins bound to the Staphylococci or Sepharose beads were removed by low speed centrifugation, and the precleared lysate was incubated with DHBV pre-S peptide (amino acids 80-102) immobilized on Sepharose beads via the GST tag (40 μl b.v. beads, circa 80 μg fusion protein) at 4°C overnight. After three times washing with the lysis buffer, the retained proteins were separated by 6% SDS-PAGE gel and blotted onto polyvinylidene difluoride (PVDF, Bio-Rad) membranes. Protein bands were revealed by staining with 0.1% Ponceau S and the 120 kDa band was cut out. An estimated 28 μg (equivalent to 20 pmol) of p120 protein was obtained from 40 gm duck liver.

Protein microsequencing of p120. The p120 blotted to PVDF membrane was sent to Harvard Microchemistry Facility for customized protein microsequencing analysis. Briefly, p120 was digested with trypsin, and digested fragments were separated by high-pressure liquid chromatography (HPLC). Selected peptides were sequenced by the Edman degradation method.

Construction of duck liver cDNA libraries. RNA was extracted from frozen Pekin duck liver with guanidinium thiocyanate (RNA isolation kit, Stratagene), and mRNA purified through an oligo dT column (Poly(A)Quik mRNA purification kit, Stratagene). Oligo dT primed and random primed lamda expression libraries were constructed using ZAP-cDNA synthesis kit and ZAP Express cDNA synthesis kit, respectively (Stratagene). For directional cloning, 1st strand cDNA synthesis was primed with oligo-dT (the primer also contains an XhoI site) using Moloney murine leukemia

virus (M-MuLV) reverse transcriptase. After 2nd strand synthesis and addition of an EcoRI adaptor, the cDNAs were digested with XhoI and size-fractionated with Sephacryl S-400 column. Fractions 4 and 5 were combined and further electrophoresed in 1% agarose gel to isolate cDNAs greater than 1 kb in size. The purified cDNAs were ligated to EcoRI/XhoI double digested Lambda ZAP II vector. The ligation product was packaged into Gigapack II Gold packaging extract (Stratagene), then plated in XL1-Blue MRF' host cells. The primary library was estimated to contain $6x10^6$ independent recombinants and has an average insert size of 1.9 kb. This oligo dT primed library was amplified once and stored in aliquots at 4°C (in chloroform) or -70°C (in DMSO). For construction of random primed library, 1st strand cDNA synthesis was primed by random hexameric oligonucleotides. The cDNAs were ligated with EcoRI adaptor, size fractionated with Sephacryl S-400 column, and cloned into EcoRI digested ZAP express vector (Stratagene). This random primed library has $6x10^6$ independent recombinants and an average insert size of 1.6 kb.

Cloning of p120 cDNA. Since amino acid sequences of all four p120 peptides perfectly matched chicken and human glycine decarboxylase gene, duck cDNA for p120 was isolated by a cDNA fragment of the chicken glycine decarboxylase. Chicken liver mRNA was reverse transcribed by random hexamers and superscript II reverse transcriptase (Gibco/BRL) at 42°C. The RNA template was removed by treatment with RNase H. The 1st strand cDNA was used as template for PCR amplification of the partial coding sequence for chicken glycine decarboxylase. The primers used were based on published sequence (Kume et al., 1991) and had the sequence 5'-ATCACTGAGCT CAAATTACCCCATGAGATG-3' (SEQ ID NO:14; sense primer, positions 679-701) and 5'-GGAAACTCGAGCTGGAAGCAGTGTTATGAA-3' (SEQ ID NO:22; antisense primer, positions 3038-3009). After 35 cycles of amplification using Vent DNA polymerase the product (2.3 kb) was purified from the agarose gel and labeled with ^{32}P-dCTP using random DNA labeling kit (Amersham). To screen for p120 cDNAs, oligo dT primed duck liver library was plated onto NZY plates at a density of $5x10^4$ pfu/150 mm plate and cultured at 37°C overnight. The plaques were transferred onto duplicate nitrocellulose filters. The filters were hybridized with the ^{32}P labeled chicken cDNA fragment, washed and exposed to X-ray films. The final wash consisted of 0.5 x

SSC/0.1% SDS at 55°C. For secondary screening, areas of NZY agar plates corresponding to positive hybridization signals were cut out and immersed in SM buffer to elute the phages. The phages were used to infect XL1 Blue MRF' cells and plated onto NZY plates at low density to prevent individual plaques from merging with each other. The duplicate filters were rescreened with the chicken cDNA probe and positive plaques identified and isolated.

p120 DNA sequencing. After the secondary screening, the pBluescript SK plasmids containing the inserts were excised from the p120 positive Lambda phages using a protocol provided by the manufacturer. DNA sequences at the ends of the inserts were determined with the aid of sequencing primers annealing to the 5' end (T3 transcription primer) or 3' end (pUC/M13 -40 primer) of pBluescript vector. For systemic sequencing of the longest insert, deletion constructs were generated using exonuclease Bal 31. To generate clones with deletions at the 5' end, the recombinant plasmid was linearized with BamHI, and treated with Bal 31. Aliquots were taken out at different time points after digestion and exonuclease activity terminated by phenol/chloroform extraction. DNA in each aliquot was digested with KpnI and cloned into KpnI/Hind II digested M13mp18. For deletions at 3' end of the insert, the recombinant plasmid was linearized with XhoI, and treated with Bal 31. DNA aliquots at different time points of exonuclease digestion were further digested with BamHI and cloned into BamHI/HincII digested M13mp18 DNA. Recombinant M13 phages were identified by DNA hybridization. Sequencing was performed on single stranded phage DNA using the universal pUC -40 sequencing primer. A few regions were not covered by the deletions and were sequenced using internal primers. Gel compressions were resolved by sequencing the opposite strand and by using dITP reagents.

Cell-free translation of p120 cDNA and recapitulation of pre-S binding capacity. The longest p120 cDNA clone, by analogy with chicken and human glycine decarboxylase, is expected to contain nearly complete coding sequence for the mature form of p120. For cell-free expression of p120 from this clone, an in-frame ATG codon surrounded by optimal Kozak sequence is attached to its 5' end via 20 cycles of PCR, using the sense primer 5'-atggtaccatgGAGGCGGCGCGGTGCATCGAGC-3' (SEQ ID NO:16) and antisense primer 5'-ATCTCGAGATATTAACATTAGCAATGTTACT-3'

(SEQ ID NO:17; small letters: nontemplated sequences). The PCR product was digested with KpnI and XhoI and cloned into the KpnI/XhoI sites of pBluescript vector. Coupled in vitro transcription/translation was carried out with T7 RNA polymerase, rabbit reticulocyte lysate and ^{35}S methionine using a kit purchased from Promega (TnT transcription/translation system). For pre-S binding assay, a fraction of translation product (5-10 μl) was incubated with 2-4 mg of various forms of DHBV pre-S protein expressed as GST fusion proteins and immobilized on Sepharose beads. The incubation was carried out in lysis buffer at 4°C for a few hrs to overnight. After vigorous wash with lysis buffer, bound p120 was revealed by SDS-PAGE followed by fluorography.

High affinity for truncated forms of pre-S protein. p120 is not a glycosylated protein since labeling primary duck hepatocytes in the presence of tunicamycin (1 μg/ml) did not modify the mobility of the binding protein. The three p120 binding constructs contained pre-S sequences 92-161, 98-161, and 1-102, respectively. These results suggest that p120 binding motif is normally hidden by the surrounding pre-S sequences but can be made accessible by substantial truncation at either N- or C- terminus. Further deletion removed p120 binding motif, abrogating p120 binding. Thus the sequence bracketed by residues 98 and 102 is the putative p120 binding motif. Consistent with this interpretation, p120 binding capacity was maintained in the three double deletion constructs with a fixed C-terminus at residue 102 but different N-termini at residues 25, 59, and 80. Construct 80-102, which retained only 23-aa residues of the pre-S region, bound p120 at least as efficiently as the longer construct 1-102. To test whether residues 98-102 can bind p120 in the absence of any surrounding DHBV sequence, mutant E98S-102 was constructed (residue 98 in this construct was converted from glutamic acid to serine due to the enzymatic manipulation). This construct, as predicted, was able to bind p120 efficiently.

The p120 binding site overlaps extensively with the binding site (residues 100-107) of a virus-neutralizing monoclonal antibody, SD20 (Chassot et al. (1993) Virology 192:217-223). This neutralizing epitope is one of the three clustered neutralizing epitopes.

We also isolated a monoclonal antibody which at a 1:200 dilution of the hybridoma culture supernatant inhibited DHBV infection in primary duck hepatocytes by

more than 90%. With the use of GST tagged deletion mutants, the binding site of this neutralizing monoclonal was mapped to pre-S residues 98-104, which entirely covers the p120 binding site. Unlike p120, the monoclonal antibody requires pre-S residues 103 and 104 for binding.

Retention of low level p120 by intact pre-S construct. The fact that intact pre-S protein failed to retain p120 raised the issue of the significance of p120-pre-S interaction. To test the possibility that intact pre-S protein can immobilize small amounts of p120, we used a more sensitive detection method. Unlabeled liver proteins from 0.5 gm tissue were incubated with the GST fused intact pre-S protein (1-161) and bound p120 protein was identified by Western blot analysis with a rabbit polyclonal antiserum raised against gel-purified p120. With this improved sensitivity detection system, p120 was found in the retained material, though much less efficiently than retention by construct 80-102. p120 retention by the intact pre-S protein seems independent of p170, since it occurs in DHBV-infected liver in which p170 binding is sequestered by endogenous viral pre-S protein.

Efficient p120 binding requires precise truncation at the pre-S C-terminus. While all constructs terminating at residue 102 (1-102, 25-102, 59-102, 80-102) bound large amounts of p120, those ending at residue 104 (1-104, 25-104, 80-104) did not (Fig. 1). To further define the boundary between p120-binding and nonbinding C-terminal deletion constructs, three additional constructs with fixed N-terminus at residue 80 but different C-termini at residues 103, 101, and 100 were constructed. p120 did not bind the mutant truncated at amino acid 104, but could bind mutants terminating at either residue 103 or 102. Further truncation to residue 101 or 100 abrogated binding. Therefore, C-terminal truncation has to occur at residue 102 (with a C-terminus Phe-Arg-Arg) or 103 (with a C-terminus Arg-Arg-Gln) in order to gain affinity for p120.

p120 binding motif is composed of a tripeptide sequence of Phe-Arg-Arg. To further elucidate the contribution of individual amino acid residues to p120 binding, site-directed mutagenesis was carried out for residues 97 through 102. Most mutants were generated in the 80-102 construct, because of its strong reactivity with p120 in wild-type sequence as well as for simplicity in the construction of mutations. The binding results of the pre-S mutants in the construct 80-102 are shown in the left panel of Fig. 23. Mutating

glutamic acid at position 98 to valine maintained, and increased, binding. Changing alanine at position 99 to aspartic acid also retained reactivity towards p120 (A99D). This residue is not conserved in nature as it was found to be changed to tyrosine in a goose hepatitis B virus strain which can infect ducks (Shi et al., GenBank Accession No. M95589). As to the phenylalanine residue at position 100, substitution by either valine or leucine abolished binding (F100V, F100L), while mutation to another aromatic amino acid such as tryptophan was tolerated (F100W). For the penultimate arginine, mutation to neither leucine nor histidine was tolerated (R101L, R101H), but substitution by lysine (R101K) retained residual binding capacity. Most stringent for p120 binding was the terminal arginine residue, for which substitutions by glycine, histidine, and even by lysine, totally abolished p120 binding (R102G, R102H, R102K). Therefore, the triplex of Phe-Arg-Arg residues at positions 100-102 are critical for p120 binding and very likely constitute the p120 contact site. These three residues are conserved in all DHBV strains (and a goose hepatitis B virus strain) sequenced.

Residues 100-102 are positioned at the C-terminus in the construct 80-102. To rule out the possibility that the importance of these residues was merely a result of positional effect, three mutants of these residues, F100V, R101L, R102G, and three mutants of residues 97 and 98, R97C, E98A and E98V, were introduced into another p120 binding construct 92-161 and the affinity for p120 was tested. All the three mutants of residues 100-102 failed to retain p120 while mutants of residues 97 and 98 could bind to the protein (Fig. 27, right panel).

p120 is detectable on the hepatocyte cell surface. If p120 is (or part of the) DHBV receptor, it should be expressed on the cell surface. The possible cell surface distribution of p120 was examined by labeling primary duck hepatocytes with sulfo-LC-biotin, which is too bulky to penetrate the surface of viable cells. Since labeling was carried out on hepatocytes attached to dishes, dead cells could be removed by the successive PBS washing prior to and following labeling. Incubation of the biotinylated lysates with several pre-S constructs revealed availability of p120 on the cell surface (Fig. 26). It was detected very efficiently by the construct 80-102 (lane 1) and weakly by construct 25-102 (lane 3), but not by constructs 80-104 (lane 2) and 1-161 (lane 4). The much stronger intensity of p120 retained by 80-102 versus 25-102 probably reflects non-

linearity of signal amplification by ECL. Under the same conditions p170 was retained by constructs 1-161 and 80-104, though the signal obtained was weaker. To rule out the possibility that the p120 and p170 detected were derived from a small amount of intracellular protein leaked from dead cells, immunoprecipitation of Golgi ß-COP, a Golgi microtubule-associated protein was performed using monoclonal antibody M3A5. This mAb recognizes an epitope shared by the Golgi ß-COP protein (110 kd) and a high MW doublet of microtubule-associated protein (MAP). Both 110 kd protein band (not shown) and a doublet high MW band (~300 kd) were precipitated only from total labeled duck liver lysate (Figure 26, lane 6), but not from surface labeled lysate (lane 5), suggesting that contamination by cytoplasmic proteins was insignificant in these experiments.

Disruption of p120 binding motif reduces DHBV infectivity. The crucial role of pre-S residues 100-102 in mediating p120 interaction enabled us to test the significance of p120 in the DHBV life cycle by genetic approaches. Double amino acid substitutions were introduced into p120-binding residues 100-102 and succeeding residues 103-104 of the replication competent DHBV genome pDHBV3.5: F100V/R101L, R101I/R102D, R101L/R102L, and Y103C/Q104F. As a control, a mutant with triple amino acid substitutions, K95S/R97L/E98A, was used (the mutations cover the p170 binding site but do not abolish p170 binding). Although some of the mutations caused amino acid changes in the overlapping polymerase gene (R101I/R102D: S311Y/P312R; F100V/R101L: F310C; Y103C/Q104F: S314F; K95S/R97L/E98A: S306C), this portion of the polymerase is a spacer region tolerant of substantial sequence alterations. After the mutants were transfected into LMH cells, secretion of pelletable particles into culture medium at different time points were measured. In general, no major variation in secretion of virion/core particles was found except that mutant Y103C/Q104F produced fewer particles at day 7 post-transfection. All the mutants displayed similar ratio of virion/core particles, since removal of core particles by Pronase/DNaseI decreased hybridization signal similarly for all the mutants (Fig. 27). No major difference in levels of cccDNA, total and core DHBV RNA, and core DNA could be detected at day 7 post-transfection.

Equal amounts of virion particles concentrated from culture medium were used to infect primary duck hepatocytes for 6.5 hr. The infectivity was measured by Southern blot hybridization of intracellular DHBV DNA at day 8 post-infection. All the mutants except the triple mutant K95S/R97L/E98A exhibited significantly reduced amount of viral DNA in infected cells. According to results from four independent transfection/infection experiments, the degree in reduction of infectivity follows the order Y103C/Q104F > F100V/R101L and R101I/R102D > R101L/R102L. Immunofluorescence staining of infected cells with an anti-pre-S antibody revealed a corresponding reduction in the number of cells infected, though the intensities of fluorescence in the positive cells were not significantly different between cells infected with wild-type virus and mutants.

Synthetic pre-S peptides covering the p170/p120 binding sites interfere with DHBV infectivity. The p170 binding site entirely covers the three clustered neutralizing epitopes, and p120 binding site overlaps the C-terminal epitope. If the clustered epitopes are part of receptor binding site, then pre-S peptides covering this region might compete for receptor binding and interfere with DHBV infection. Two pre-S peptides were used for this experiment: 80-102 and 80-104. The GST fusion protein of peptide 80-102 binds p120 (but not p170) efficiently, while peptide 80-104 binds p170 (but not p120) with low efficiency. The two pre-S peptides were purified from GST by thrombin cleavage. Peptides were pre-incubated with hepatocyte monolayers for 1 hr before infection with 1 µl of viremic duck serum. As a result, both peptides reduced DHBV infectivity. At the concentration of 10 mg/ml, peptides 80-102 and 80-104 reduced DHBV infectivity to similar degrees. Increasing the peptide concentration to 1 mg/ml enhanced inhibitory effect significantly for peptide 80-104 but only slightly for peptide 80-102. The strong inhibitory effect of peptide 80-104 lends support to the hypothesis that the clustered neutralizing epitopes are a contact site of DHBV receptor.

p120 expression is restricted to DHBV infectable tissues. To study the tissue distribution of p120, tissue lysates were precleared and incubated with the construct 80-102, and retained proteins revealed by SDS-PAGE minigel followed by Coomassie blue staining. p120 was clearly found in the liver and kidney (Fig. 29). It was weakly detected in pancreas but not in other tissues examined, including stomach, lung, small intestine, skeletal muscle, and spleen, heart, gall bladder. As an independent

confirmation of this result, a direct Western blot detection method for p120 was developed. In accordance with the affinity approach, p120 was found most abundantly in the liver followed by kidney, but weakly in pancreas (Fig. 29). No p120 was found in other tissues tested. This pattern of tissue specific distribution of p120 coincides with the known tissue tropism of DHBV infection.

Four peptide sequences of p120 were obtained (Fig. 30). Data base search revealed complete homology of all the peptides with chicken and human glycine decarboxylase. This result suggests that the p120 pre-S binding protein is actually the duck form of glycine decarboxylase. Based on this assumption, p120 cDNAs were cloned from duck liver cDNA libraries by cross-hybridization with a 2.3 kb fragment of the chicken glycine decarboxylase cDNA. Several cDNA molecules were obtained from oligo-dT primed library. The longest clone was studied in detail. The complete nucleotide and deduced amino acid sequences of the coding sequence of this clone are shown in Fig. 31. The clone contains 2922 nucleotides encoding 973 amino acid residues. All four peptide sequences obtained from microsequencing could be found within this coding region. Compared to chicken glycine decarboxylase, the sequence homology at amino acid level is 98% (950/973) and there is a 6-amino acid insertion at the N-terminal region of the duck protein. To test whether the duck glycine decarboxylase is indeed the pre-S binding protein we characterized, the coding region of this clone was subcloned into pBluescript vector and encoded protein translated in vitro in rabbit reticulocyte lysate. Although this clone lacks 37 N-terminal amino acid residues compared with the putative precursor protein of the chicken gene, it misses only three amino acid residues as compared with the mature form of the chicken protein. The in vitro translated protein had molecular mass about 120 kD, close to the size observed for the pre-S binding protein detected in duck liver. When incubated with various forms of GST-DHBV pre-S proteins, the radiolabeled protein could be retained only by pre-S protein with either C-terminal truncation at a 102 or N-terminal truncation at 92 (Fig. 24, lane 2, 4). It could not be retained by the intact pre-S protein nor by a construct with C-terminal truncation at 104 (Fig. 24, lane 1, 3). Moreover, single point mutation in residue 101 or 102, which form part of p120 contact site, abolished retention of the in vitro translated duck glycine decarboxylase protein (Fig. 24, lane 5, 6). This pattern of

selective binding to a few truncated forms of pre-S protein through a critical contact site is in complete accordance with results obtained with p120 from duck liver, confirming that the p120 pre-S binding protein is indeed glycine decarboxylase.

Properties of p120. Compared with p170, p120 exhibits remarkable tissue specific expression in the liver and the kidney, the major sites of DHBV replication. The data indicating that p120 is a candidate DHBV binding protein of a putative receptor complex is supported by: 1) restricted tissue distribution coincident with the tissue tropism of DHBV replication; 2) co-localization of the binding site with virus-neutralizing epitopes; 3) interference of DHBV infection by short pre-S peptides covering its (and p170's) binding site; 4) reduced infectivity of DHBV mutants with a disrupted p120 binding motif; 5) cell surface localization on primary duck hepatocytes.

p120 is not exclusively a cell-surface protein. Although we could detect p120 by cell surface labeling with sulfo-LC-biotin, the protein is also located within the hepatocytes. Whether the amount of p120 present on cell surface is sufficient to allow viral entry is not presently known. Alternatively, p120 could play a role in the intracellular trafficking of internalized DHBV particles and if this proves to be the case, it does not need to be located principally on the cell surface. It was determined that a homologous binding protein was detectable in DHBV non-infectable hosts. Using the pre-S construct 80-102, we could detect p120-related proteins in chicken and human liver. If the host and tissue specificities of DHBV infection are determined at the receptor level, the double specificities may be explained by postulating that p170 and p120 proteins are part of a DHBV receptor complex.

The p120 reactive peptide 80-102 does not inhibit DHBV replication to as great an extent as p170 reactive peptide 80-104. A possible explanation for this phenomenon is that DHBV-hepatocyte interaction is initiated by binding to p170, followed by binding to p120. A peptide that inhibits the first stage of interaction (80-104) would potentially be a more potent inhibitor of infection than a peptide that inhibits a later stage of this interaction (80-102). As for the observation that the inhibitory effect is dose-dependent for 80-104 but not obviously for 80-102, a difference in their affinity for the respective target protein is one possible explanation. In this regard, 80-102 has a strong affinity for p120 and its maximum inhibitory effect is probably reached at the lowest peptide

concentration used; whereas 80-104 has low affinity for p170 and its maximal inhibitory effect requires the highest peptide concentration.

As shown in Fig. 23, intact pre-S protein expressed in E. coli retains only a small amount of p120. Addition of DHBV viremic duck serum into a tube containing metabolically labeled hepatocyte lysate and immobilized pre-S construct 80-102 failed to significantly reduce the amount of p120 retained to the beads, suggesting DHBV particles do not have large amount of truncated pre-S protein similar to those seen in vitro to bind p120 efficiently. However, considering the fact that residues 101 and 102 are both arginine, p120-reactive pre-S protein can be generated by proteolytic cleavage through a trypsin-like protease or an endopeptidase specific for dibasic residues (see Davey et al. (1994) EMBO J. 13:5910-5921). Proteolytic cleavage of viral envelope protein after di- or tetra-basic residues is required for infectivity of myxoviruses and retroviruses, although in these instances the cleavage: 1) occurs during virion maturation; and 2) exposes a new hydrophobic N-terminus required for virus-cell fusion. In addition, Lu and colleagues recently demonstrated that protease treatment of HBV particles enhanced its infectivity in a hepatoma cell line (Lu et al. (1996) J. Virol. 70:2277-2285), thus reinforcing the hypothesis that hepadnavirus infection may require the action of proteases.

The simplest model to account for p170 and p120 as components of DHBV receptor is as follows: virion particles are attracted onto the hepatocyte surface initially by p170. This event is followed by either direct conformational change of the pre-S protein or by proteolytic cleavage at Arg102 to activate p120 binding. p120 binding subsequently allows viral entry or participates in intracellular trafficking. Irrespective of how the p120-DHBV interaction occurs, the role of p120 in DHBV life cycle is testable by two experimental approaches: namely, whether antibody against p120 inhibits DHBV infection of primary duck hepatocytes, and/or whether transfection of p120 cDNA (together with p170 cDNA) into nonpermissive cells renders them susceptible to DHBV infection.

Activity Assays. Each such homolog can be definitively identified as a hepadnaviral receptor by any of the following assays:

1) Antibody inhibition experiment. Primary human hepatocytes can be obtained by perfusion, cultured, and infected with hepadnavirus according to the method of Gripon et al. (Virology, 192: 534-540, 1993) and cultured. An antibody, e.g., a rabbit polyclonal antibody, is added to the culture medium during the stage when virus particles are put on cell monolayer, preferably in a 1:100 to 1:1000 dilution. After infection, cells are maintained for a week. Cellular DNA analyzed by Southern blot for the presence of HBV DNA. If the addition of antibody to the culture results in a reduced level, or complete absence, of HBV compared with the amount of HBV DNA detected in control cells infected in the absence of receptor-specific antibody, then the antibody blocks viral infectivity, and the receptor to which it specifically binds is a hepadnavirus receptor.

2) Cell Transfection Assay. To assay the ability of a hepadnavirus receptor cDNA to render a cell line susceptible to HBV infection, the cDNA is first transfected into the cells. For this purpose, the coding sequence of the receptor gene is cloned into a vector for expression in mammalian cells. An appropriate vector is one that contains a selectable marker, e.g., neomycin resistance, a variety of appropriate vectors being commercially available. Methods of transfecting cloned DNA into an established cell line is a basic technique known to those of ordinary skill in the art. The usual method is by $CaCl_2$ precipitation. After transfection, a reagent such as neomycin is applied to select for cells receiving the plasmid DNA containing the gene of interest. These cells are cloned and used for further analysis. The second step is to test whether these transfected cell lines are newly infectable with HBV. No known cell lines are susceptible to HBV infection. However, a few human hepatoma cell lines support HBV replication when HBV DNA is transfected into them. This suggests that while their intracellular machinery is compatible for the HBV life cycle, they do not allow HBV adsorption or penetration, with a defect probably at the receptor level. These cell lines include, e.g., HepG2 cells and Huh7 cells. Thus, these cell lines should be the first choice for transfection by the cDNA of interest. To infect the cells, HBV particles are transfected into the cell line of choice. The cells are incubated for a few hours to overnight and then thoroughly washed to remove any unbound virus. After further culture for one to two weeks, cells are lysed. Cellular DNA is tested by Southern blot analysis for the presence of viral DNA. The virus stock could come from serum taken from hepatitis B patients

positive for HBeAg, or it could be concentrated from medium of hepatoma cells transfected with cloned HBV DNA. Evidence of viral replication unambiguously demonstrates that the gene of interest encodes the cellular receptor specific for HBV.

Preparation of the isolated DNAs of the invention

Some alternative means of preparing the nucleic acids of the invention, using the information provided herein and standard techniques, are as follows:

(1) A nucleic acid fragment having a nucleotide sequence shown in Figs. 18-19, or a nucleic acid encoding the amino acid sequence shown in Figs. 18-19, but, owing to the degeneracy of the genetic code, having a nucleotide sequence different from that shown in the figure, may be synthesized by standard chemical means as generally applied to synthesis of oligonucleotides.

(2) An isolated DNA prepared by any of the methods outlined herein (including the methods originally used to obtain the DNAs of the invention) may be used to probe an appropriate cDNA library or genomic DNA library. Preferably, a human liver cDNA library is used, either oligo-dT primed or primed by random hexamer primers. The ideal library is derived form normal human liver rather than hepatoma or cultured hepatocyte lines, because hepatoma or hepatocyte lines are generally not infectable with HBV. The library is constructed in a lamda vector as a bacteriophage.

To screen the library, phages are grown in NZY medium at around $5x10^4$ per plate, transferred to a nitrocellulose filter, and hybridized with a nucleic acid of the invention, e.g., the 2.5 kb p170 cDNA randomly labeled by ^{32}P dCTP. The blots are then washed at approximately 50°C with 2XSSC/0.1% SDS. Positive clones are sequenced to verify homology to p170. The stringency of the washing step after hybridization can be adjusted from low to high, and the clone that hybridizes with p170 at the highest stringency is most likely to be a homolog, e.g., a human homolog, of p170. High stringency hybridization can be performed by hybridizing in 50% deionized formamide, 800 mM NaCl, 20 mM Pipes, pH 6.5, 0.4% SDS, 500 μg/ml denatured, sonicated salmon sperm DNA at 42°C for 12-20 hours; and washing in 30 mM NaCl, 3.0 mM sodium citrate, 0.5% SDS at 65°C). It is expected that hybridization and wash conditions such as the highly stringent conditions set forth above would be adequate; if necessary, the stringency may be increased or decreased, without undue experimentation, using methods

well known to those of ordinary skill in the art (see, e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989). The stringency of washing is dependent on two factors primarily: temperature (the higher the more stringent) and SSC concentration (the lower the more stringent). The identity of the homolog is verified by its binding affinity for the pre-S domain of the appropriate hepadnavirus, e.g., the human HBV pre-S domain. Affinity can be determined by transfecting the cDNA into a cell line which normally does not bind the human HBV particle, labeling the cell with ^{35}S methionine, and reacting the cell lysate with a GST fusion protein of the human HBV pre-S protein, as was shown above for the DHBV pre-S-p170 interaction.

Anti-hepadnavirus receptor antibodies. Cell receptor or receptor fragments of the invention may be used to generate antibodies by conventional methods well known to those skilled in the art, including those which generate polyclonal antibodies and those which generate monoclonal antibodies (see, e.g., Coligan et al., eds. Current Protocols in Immunology, Wiley & Sons, Inc., 1994). For example, the deduced amino acid sequence of the p170 receptor can be used to guide the selection of regions of the receptor protein which would be likely to be exposed on the cell surface, and thus would be presented to antibodies in vivo. A short peptide representing one or more of such regions may be synthesized (e.g., chemically or by recombinant DNA techniques) and used to immunize an animal (e.g., a rabbit or a mouse) to generate polyclonal or monoclonal antibodies. For example, certain of the peptides 1-4 shown in Fig. 14A can be chemically synthesized using standard techniques. Alternatively, the p170 amino acid sequence responsible for binding the pre-S domain should be virus neutralizing and thus would be a suitable peptide antigen. The peptides are used to generate polyclonal antibodies in rabbits by the following procedure:

A preparation of a given peptide emulsified with complete Freund's Adjuvant is injected intradermally into rabbits. Booster injections are emulsified in complete adjuvant and injected at monthly intervals.

Antibody titer is assessed using either of two methods. First, serial dilutions of the antiserum in 1% normal rabbit serum are incubated with ^{125}I-labelled p170 receptor fragment by standard methods (e.g., see Segre et al., J. Biol. Chem. 254:6980, 1979) for

24 h at 4°C. The bound ^{125}I-p170 receptor fragments are separated from unbound fragments by addition of 100 µl of second antibody (anti-rabbit IgG, Sigma) diluted 1:20 and 1 ml of 5% polyethylene glycol, followed by centrifugation at 2000 rpm for 30 min. at 4°C. The supernatant is removed and the pellet analyzed for radioactivity in a γ-counter. In the second method, cell lines expressing recombinant hepadnavirus receptor (COS cells or CHO cells transfected with a nucleic acid encoding a hepadnavirus receptor) are incubated with serially diluted antibody at 4°C, 20°C or 37°C for 1- 4 h. The cells are rinsed with PBS (x3) and incubated for 2 h at 4°C with ^{125}I-labelled (NEN, Dupont) or FITC-labelled (Sigma) second antibodies. After rinsing (x3 with PBS), the cells were either lysed with 0.1 M NaOH and counted in γ-counter (if ^{125}I-labelled second antibody was used) or fixed with 1% paraformaldehyde and examined by fluorescent microscopy (if FITC-labelled second antibody was used).

Another method for producing antibodies utilizes as antigen the intact cell receptor protein of the invention expressed on the surface of cells (e.g., mammalian cells, such as COS cells, transfected with DNA encoding the receptor). Such cells are prepared by standard techniques, e.g., by the DEAE-dextran transfection method, using a vector encoding and capable of directing high-level expression of the cell receptor. Such cells may be used to generate polyclonal or monoclonal antibodies. For example, monoclonal antibodies specific for the hepadnavirus receptor can be produced by the following procedure:

Intact COS cells expressing high levels of hepadnavirus receptor on the cell surface are injected intraperitoneally (IP) into Balb-c mice (Charles River Laboratories, Wilmington, MA). The mice are boosted every 4 weeks by IP injection, and are hyperimmunized by an intravenous (IV) booster 3 days before fusion. Spleen cells from the mice are isolated and are fused by standard methods to myeloma cells. Those hybridomas which produce antibodies capable of binding to the hepadnavirus receptor are cultured and subcloned.

Mapping the pre-S binding site on the hepadnavirus receptor. Several experimental approaches can be envisaged for mapping the binding site on the receptor.

(1) The nucleic acid encoding the receptor is cloned into a transcription vector, transcribed, and translated in rabbit reticulocytes in the presence of radioactive tracer

(35S methionine). The labeled lysate is incubated with a GST-pre-S fusion protein. If binding is positive and specific, a series of deletion mutants of the receptor gene can be made, transcribed, translated, and reacted with GST-pre-S fusion protein.

(2) The nucleic acid encoding the receptor, or a fragment thereof, is cloned into the pGEX 2TK vector (Pharmacia) and expressed as a GST fusion protein. A hepadnavirus pre-S protein is expressed in eukaryotic cells by stable transfection-selection (as described above). The pre-S expressing cells are labeled and lysate is reacted with a GST fusion protein of the hepadnavirus receptor gene to determine the specificity of the interaction.

(3) The nucleic acid encoding the receptor gene is expressed in eukaryotic cells and interacted with a GST pre-S fusion protein. This is the reciprocal of (2), above. This is the preferred approach if posttranslational modifications such as glycosylation or phosphorylation are essential for binding of the hepadnavirus receptor to pre-S protein.

Expression of Polypeptides. Polypeptides according to the invention may be produced by expression from a recombinant nucleic acid having a sequence encoding part or all of a cell receptor of the invention, using any appropriate expression system: e.g., transformation of a suitable host cell (either prokaryotic or eukaryotic) with the recombinant nucleic acid in a suitable expression vehicle (e.g., pCDNA I). The precise host cell used is not critical to the invention; however, the following host cells are preferred: COS cells, CHO cells, and human liver cells. Mammalian cell transfection methods are described, e.g., in Ausubel et al. (Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989); expression vehicles may be chosen from those discussed, e.g., in Cloning Vectors: A Laboratory Manual (P.H. Pouwels et al., 1985, Supp. 1987). Stably transfected cells are produced via integration of receptor DNA into the host cell chromosomes. Suitable DNAs are inserted into pcDNA, pCDNA I-Neo, or another suitable plasmid, and then cells are transfected with this plasmid with or without cotransfection with psV-2-Neo, or psV-2-DHFR by standard electroporation, calcium phosphate, and/or DEAE/Dextran techniques. Selection of transfected cells is performed using progressively increasing levels of G418 (Geneticin, GIBCO), and if necessary, methotrexate.

DNA sequences encoding the polypeptides of the invention can also be expressed in a prokaryotic host cell. DNA encoding a cell receptor or receptor fragment is carried on a vector operably linked to control signals capable of effecting expression in the prokaryotic host. If desired, the coding sequence may contain, at its 5' end, a sequence encoding any of the known signal sequences capable of effecting secretion of the expressed protein into the periplasmic space of the host cell, thereby facilitating recovery of the protein and subsequent purification. Prokaryotes most frequently used are various strains of E. coli; however, other microbial strains may also be used. Plasmid vectors are used which contain replication origins, selectable markers, and control sequences derived from a species compatible with the microbial host. For example, E. coli may be transformed using derivatives of pBR322, a plasmid constructed by Bolivar et al. (Gene 2: 95, 1977) using fragments derived from three naturally-occurring plasmids, two isolated from species of Salmonella, and one isolated from E. coli. pBR322 contains genes from ampicillin and tetracycline resistance, and thus provides multiple selectable markers which can be either retained or destroyed in constructing the desired expression vector. Commonly used prokaryotic control sequences (also referred to as "regulatory elements") are defined herein to include promoters for transcription initiation, optionally with an operator, along with ribosome binding site sequences. Promoters commonly used to direct protein expression include the beta-lactamase (penicillinase), the lactose (lac) (Chang et al., Nature 198: 1056, 1977) and the tryptophan (Trp) promoter systems (Goeddel et al., Nucl. Acids Res. 8: 4057, 1980) as well as the lambda-derived PL promoter and N-gene ribosome binding site (Simatake et al., Nature 292:128, 1981).

The nature of the cell receptor proteins of the invention is such that, upon expression within a cell, it is moved to the cellular membrane and partially through the membrane, so that part of it remains embedded in the membrane, part extends outside the cell, and part remains within the cell. Transformed cells bearing such embedded cell receptors may themselves be employed in the methods of the invention, or the receptor protein may be extracted from the membranes and purified.

Expression of peptide fragments lacking the hydrophobic portions of the protein responsible for anchoring the intact protein in the cellular membrane would not be expected to become embedded in the membrane; whether they remain within the cell or

are secreted into the extracellular medium depends upon whether or not a mechanism promoting secretion (e.g., a signal peptide) is included. If secreted, the polypeptide of the invention can be harvested from the medium; if not, the cells must be broken open and the desired polypeptide isolated from the entire contents of the cells.

The polypeptide of the invention can be readily purified using affinity chromatography. Antibodies to these polypeptides, or the receptor specific ligands (e.g., a pre-S polypeptide) may be covalently coupled to a solid phase support such as Sepharose 4 CNBr-activated sepharose (Pharmacia), and used to separate the polypeptide of the invention from any contaminating substances. Typically 1 mg of ligand or antibody will be incubated with CNBr-activated sepharose at 4°C for 17-20 h (with shaking). The sepharose is rinsed with 1 M Tris HCl (pH8) to block excess active sites. The Sepharose-p170, Sepharose-pre-S, or Sepharose-antibody is then incubated with the crude polypeptide in phosphate-buffered saline (pH 7.4) at 4°C for 2 h (with shaking). The Sepharose is then typically packed in a column, thoroughly washed with PBS (typically 10 times the column volume), and eluted with dilute HCl in H_2O (pH 1.85). The eluate may then be concentrated by lyophilization and its purity checked, for example, by reverse phase HPLC.

Screening for hepadnavirus receptor antagonists and agonists. Candidate antagonists and agonists may be screened for the ability to compete with or enhance binding of the pre-S domain to the hepadnavirus receptor using the assays described herein.

In one example, those antibodies that recognize the hepadnavirus receptor on the intact cells are screened for their ability to compete with a form of hepadnavirus envelope protein, e.g., a pre-S polypeptide, for binding to a hepadnavirus receptor. Cells expressing hepadnavirus receptor on the cell surface are incubated with the ^{125}I-pre-S analog in the presence or absence of the polyclonal or monoclonal antibody to be tested for 4 h at 15°C. The antibody used may be from crude antiserum, cell medium, or ascites, or in purified form. After incubation, the cells are rinsed with binding buffer (e.g., physiological saline), lysed, and quantitatively analyzed for radioactivity using a gamma-counter. Antibodies that reduce binding of the pre-S analog to the hepadnavirus receptor are classified as competitive; those which do not are noncompetitive.

Therapeutic Use

Therapeutic administration of a mutant polypeptide can be accomplished using the polypeptide directly or by administering the polypeptide with gene therapy techniques. A nucleic acid that included a promoter operatively linked to a sequence encoding a polypeptide of the invention is used to generate high-level expression of the polypeptide in cells transfected with the nucleic acid. Gene transfer can be performed ex vivo or in vivo. To administer the nucleic acid ex vivo, cells can be removed from the body of the patient, transfected with the nucleic acid encoding the mutant polypeptide, and returned to the patient's body. Alternatively the nucleic acid can be administered in vivo, by transfecting the nucleic acid into target cells (e.g., hepatocytes) so that the mutant polypeptide is expressed in situ.

The nucleic acid molecule is contained within a non-replicating linear or circular DNA or RNA molecule, or within an autonomously replicating plasmid or viral vector, or may be integrated into the host genome. Any vector that can transfect a cell can be used in the methods of the invention. Preferred vectors are viral vectors, including those derived from replication-defective hepatitis virus (e.g., HBV and HCV), retrovirus (see, e.g., WO89/07136; Rosenberg et al., N. Eng. J. Med. 323(9):570-578, 1990; Miller et al., 1993 supra), adenovirus (see, e.g., Morsey et al., J. Cell. Biochem., Supp. 17E, 1993; Graham et al., in Murray, ed., Methods in Molecular Biology: Gene Transfer and Expression Protocols. Vol. 7, Clifton, NJ: the Human Press 1991: 109-128), adeno-associated virus (Kotin et al., Proc. Natl. Acad. Sci. USA 87:2211-2215, 1990), replication defective herpes simplex virus (HSV; Lu et al., Abstract, page 66, Abstracts of the Meeting on Gene Therapy, Sept. 22-26, 1992, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), and any modified versions of these vectors. Other preferred viral vectors include those modified to target a specific cell type (see, e.g., Kan et al. WO 93/25234; Kasahara et al. Science, 266:1373-76, 1994; Dornburg et al. WO 94/12626; Russell et al. WO 94/06920). Methods for constructing expression vectors are well known in the art (see, e.g., Molecular Cloning: A Laboratory Manual, Sambrook et al., eds., Cold Spring Harbor Laboratory, 2nd Edition, Cold Spring Harbor, New York, 1989).

Appropriate regulatory sequences can be inserted into the vectors of the invention using methods known to those skilled in the art, e.g., by homologous recombination (Graham et al., J. Gen. Virol. 36:59-72, 1977), or by other appropriate methods (Sambrook et al., eds., supra). Promoters are inserted into the vectors so that they are operatively linked 5' to the nucleic acid sequence encoding the mutant polypeptide. Any promoter that is able to initiate transcription in a target cell can be used in the invention. For example, non-tissue specific promoters, such as the cytomegalovirus (DeBernardi et al., Proc. Natl. Acad. Sci. USA 88:9257-9261, 1991, and references therein), mouse metallothionine I gene (Hammer, et al., J. Mol. Appl. Gen. 1:273-288, 1982), HSV thymidine kinase (McKnight, Cell, 31:355-365, 1982), and SV40 early (Benoist et al., Nature, 290:304-310, 1981) promoters may be used. Preferred promoters for use in the invention are hepatocyte-specific promoters, the use of which ensures that the mutant polypeptides are expressed primarily in hepatocytes. Preferred hepatocyte-specific promoters include, but are not limited to the albumin, alpha-fetoprotein, alpha-1-antitrypsin, retinol-binding protein, and asialoglycoprotein receptor promoters. Additional viral promoters and enhancers, such as those from herpes simplex virus (types I and II), hepatitis virus (Types A, B, and C), and Rous sarcoma virus (RSV; Fang et al., Hepatology 10:781-787, 1989), can also be used in the invention.

The mutant polypeptides of the invention, and the recombinant vectors containing nucleic acid sequences encoding them, may be used in therapeutic compositions for preventing or treating HBV infection. The therapeutic compositions of the invention may be used alone or in admixture, or in chemical combination, with one or more materials, including other mutant polypeptides or recombinant vectors, materials that increase the biological stability of the oligonucleotides or the recombinant vectors, or materials that increase the ability of the therapeutic compositions to penetrate hepatocytes selectively. The therapeutic compositions of the invention can be administered in pharmaceutically acceptable carriers (e.g., physiological saline), which are selected on the basis of the mode and route of administration, and standard pharmaceutical practice. Suitable pharmaceutical carriers, as well as pharmaceutical necessities for use in pharmaceutical formulations, are described in Remington's Pharmaceutical Sciences, a standard reference text in this field.

The therapeutic compositions of the invention can be administered in dosages determined to be appropriate by one skilled in the art. An appropriate dosage is one which effects a reduction in a disease caused by HBV infection. It is expected that the dosages will vary, depending upon the pharmacokinetic and pharmacodynamic characteristics of the particular agent, and its mode and route of administration, as well as the age, weight, and health (including renal and hepatic function) of the recipient; the nature and extent of the disease; the frequency and duration of the treatment; the type of, if any, concurrent therapy; and the desired effect. It is expected that a useful dosage contains between about 0.1 to 100 mg of active ingredient per kilogram of body weight. Ordinarily a dosage of 0.5 to 50 mg, and preferably, 1 to 10 mg of active ingredient per kilogram of body weight per day given in divided doses, or in sustained release form, is appropriate.

The therapeutic compositions of the invention may be administered to a patient by any appropriate mode, e.g., parenterally, as determined by one skilled in the art. Alternatively, it may by necessary to administer the treatment surgically to the target tissue. The treatments of the invention may be repeated as needed, as determined by one skilled in the art.

The invention also includes any other methods which accomplish in vivo transfer of nucleic acids into target cells. For example, the nucleic acids may be packaged into liposomes, non-viral nucleic acid-based vectors, erythrocyte ghosts, or microspheres (microparticles; see, e.g., U.S. Patent No. 4,789,734; U.S. Patent No. 4,925,673; U.S. Patent No. 3,625,214; Gregoriadis, Drug Carriers in Biology and Medicine, pp. 287-341 (Academic Press, 1979)). Further, delivery of mutant polypeptides be accomplished by direct injection of their nucleic acid coding sequences into target tissues, for example, in a calcium phosphate precipitate or coupled with lipids, or as "naked DNA".

Mutant core polypeptides and core-surface fusion proteins of the invention can be tested for their ability to inhibit hepadnavirus replication in an animal model. For example, candidate polypeptides can be injected into an animal that is infected with a hepadnavirus, e.g., a woodchuck, duck, or ground squirrel harboring its respective hepatitis B virus variants (see, e.g., Mason et al., J. Virol. 36:829, 1980; Schodel et al., in Molecular Biology of hepatitis B virus, CRC press, Boca Raton, p. 53-80, 1991;

Summers et al., Proc. Natl. Acad. Sci. USA, 75:4533-4537, 1978). Candidate polypeptides can also be analyzed in transgenic animal strains developed for the purpose of studying hepadnaviral gene expression (see, e.g., Babinet et al., Science, 230:1160-63, 1985; Burk et al., J. Virol. 62:649-54, 1988; Chisari et al., Science 230:1157-60, 1985; Chisari, in Current Topics in Microbiology and Immunology, p. 85-101, 1991). Candidate polypeptides of the invention can also be tested in animals that are naturally infected with HBV, e.g., in chimpanzees, by administering the polypeptides, or the nucleic acids encoding them, to the animal by one of the methods discussed above, or by other standard methods.

Other Embodiments

From the above description, one skilled in the art can easily ascertain the essential characteristics of the present invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. All publications cited herein are fully incorporated by reference herein in their entirety.

Other embodiments are within the claims set forth below.

What is claimed is: